プロファイリング・自動意思決定とプライバシーに関するEU（GDPR）・英国の動向

2018年9月25日

国際社会経済研究所
小泉　雄介

１．プロファイリング・自動意思決定とプライバシー等に関する論点

プロファイリングとは何か（１/２）

○ビッグデータ・AI分析におけるプロファイリングとは

- 犯罪捜査、犯罪心理学におけるプロファイリングと類似
  - ※犯罪捜査におけるプロファイリング： 犯行現場、被害者の状況、犯人の行動パターン、遺留品や遺体の状況などから、犯人の年齢、性別、職業、家族構成、性格、精神疾患の有無などの属性を推定すること。

- ある人物について、既知の情報から、その人物の既知でない情報（※1）を推定したり、将来の行動やリスク（※2）を予測すること。
  - ※1：趣味・嗜好、収入（返済能力）、健康状態・持病、性別、年代、人種・宗教等。
  - ※2：購買行動、業務パフォーマンス・適性、疾病リスク、犯罪リスク等

○プライバシー等の課題の例

- ビッグデータ・AIによって推測精度が上がることにより、推測したデータは限りなく当人の個人データと同等なものになる。
  - これにより、本人から直接取得できないデータ、本人が開示したくない属性・趣味・健康状態・年収等にについても、「個人データの取得」と実質的に同等な推測が可能となる。要配慮個人情報を推測する場合は、プライバシー侵害にもつながりうる。
- 機械学習時の入力データやアルゴリズム自体にバイアスがある場合、当該AIを用いたプロファイリングによって個人に対する社会的差別を助長したり、新たに生み出してしまう。 etc.

プロファイリングとは何か（2/2）

〇プロファイリングの例

- Amazonのレコメンド機能
  - Amazonサイトでの閲覧履歴や購買履歴に基づき、トップページでは「おすすめ商品」「チェックした商品の関連商品」「あなたのお買い物傾向から」等が表示される。また、各商品のページでは「この商品を買った人はこんな商品も買っています」「この商品をチェックした人はこんな商品もチェックしています」等が表示される。
- 行動ターゲッティング広告
  - アドネットワークではサードパーティクッキー等を用いて、利用者のWeb上での閲覧行動や広告クリック履歴等を把握しており、こうした利用者の行動履歴に合せたインターネット広告を表示している。
- Target社
  - 米国のスーパーマーケットTargetが或る10代女性の購買履歴から「妊娠している可能性が高い」とプロファイリングし、自宅に（家族がそれを知る前に）妊娠に関連する広告が送られてしまった
- Facebookの「いいね！」を用いた研究
  - ある研究では、Facebookの「いいね！」の履歴と他の情報を組合せることで、男性利用者の性的指向の88%、利用者の民族的素性の95%、利用者がキリスト教徒かイスラム教徒かの82%を正確に推測することができた。

自動意思決定とは何か

○自動意思決定」（Automated decision-making）

- EU一般データ保護規則（GDPR）（の第22条）で主に規制対象となっているのは、プロファイリング自体ではなく、プロファイリング等の自動処理のみに基づく自動意思決定である（正確にはそのうち個人に法的効果または重大な影響を及ぼすもの）。
- 例えば個人に対する信用スコア等のプロファイリングに基づいてローン審査を自動化するようなケースが、この「自動意思決定」に該当する。

○自動意思決定とプロファイリングの違い

- GDPRのプロファイリングガイドラインでは、車のスピード違反の例が挙げられている。
  - スピードカメラでの測定のみに基づいて罰金額を決める場合：
    →プロファイリングを伴わない（自動処理のみに基づく）自動意思決定
  - 今後の想定事例として、個人の運転習慣が長期に渡ってモニターされ、「常習的なスピード違反か」「直近に他の交通違反を起こしていないか」といった要素に基づいて罰金額を決める場合：
    →プロファイリング（を伴う自動処理のみ）に基づく自動意思決定
- このように「プロファイリング」と「自動意思決定」の概念は、一部重なる場合もあるが、異なる概念とされる。GDPRなど個人データ保護法制の文脈では、以下の2段階を峻別して捉える必要がある。
  - (1) プロファイリング（などの自動処理）
  - (2)（プロファイリングなどの自動処理に基づく）自動意思決定

プロファイリング・自動意思決定とプライバシー等に関する主な論点

①プロファイリングの結果、既知の個人情報から要配慮個人情報など本人の望まない個人情報が推測され、個人に紐付けられてしまったらどうするか。

- ビッグデータ・AIによって推測精度が上がることにより、プロファイリングで推測したデータは限りなく当人の個人データと同等なものになる。これにより、本人から直接取得できないデータ、本人が開示したくない属性・趣味・健康状態・年収等についても、「個人データの取得」と実質的に同等な推測が可能となってしまう。
- Webサイトによっては顧客登録画面でやたら沢山の個人情報を求めてくるサイトがあるが、そうしたサイトで自分が開示したくない情報項目を入力しなかったとしても、プロファイリングによって高い精度で推測されてしまえば、個人の（開示しないという）選択・意向は無意味なものになってしまう。特に、要配慮個人情報が推測される場合は、プライバシー侵害にもつながりうる。
- GDPRでは、通常の個人データから特別な種類の個人データ（要配慮個人情報に相当）を推測した場合も、特別な種類の個人データの取得等として取り扱わなければならないため、本人の明示的同意といった措置が必要になる。
- 日本の個人情報保護法では、通常の個人情報から要配慮個人情報を推測することは、要配慮個人情報の取得には当たらないとする見解が有力である。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（1）AIアルゴリズムや機械学習用データにおける隠れたバイアス

- 欧州データ保護監察官（EDPS）の2016年のレポートでは、AIに関する重要な懸念の1つとして、「AIの機械学習で入力に使われるデータセットに起因するバイアス」が挙げられている。AIは自身に入力されたデータのみから学習を行うが、それらのデータを大きな見地から比較考量する手段がないため、学習用データにどんなバイアスが含まれていたとしても、自らそのバイアスを回避することができない。それによって何らかの社会的差別が助長されることが懸念されている。学習用データにバイアスが無かったとしても、そもそもアルゴリズム自体に設計段階で隠れたバイアスが入っている可能性もある。

（2）一定集団への「全体責任的」なプロファイル適用

- また、一定の（被差別集団か否かに関わらず）社会集団を分類軸としたビッグデータ分析を行った場合、たとえ学習用データにバイアスが無かったとしても、当該集団に不利となる相関関係（例えば、ある人種の人々やある商品を購入した人は、ローン滞納が多い・信用リスクが高い等）が見出されることによって、その集団に属する個人にも（当人がどのような人物かに関わらず）そのようなプロファイルが当てはめられて、まるで「共同責任」であるかのように、不利益を被ってしまう（差別が再生産される）という問題も欧米で指摘されている。これは、後述⑤の予測データの正確性の問題とも関連する。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

③アルゴリズムの不透明性（ブラックボックス化）に対する説明責任をどう考えるか。

- EDPSの同じレポートでは、プライバシーや個人情報保護の上で最も重要なものの1つは「個人への透明性」であるとし、企業が個人に対してプロファイリング等に基づいて何らかの意思決定を行う場合には、当該処理について本人に情報提供しなければならないとする。しかし、そうした個人の傾向性に対する予測（や個人情報の推定）が機械学習によるAIアルゴリズムに基づいてなされる場合は、機械による理由付けの背後にあるロジックを人間の言葉で表現することは難しいとしている。これは、いわゆるAIのブラックボックス化と呼ばれる問題である。
- GDPRでは（プロファイリングなどの自動処理に基づく自動意思決定を行い、それが本人に法的効果や同様に重大な影響をもたらす場合には）、本人に対して「関連するロジックについて意味のある情報」と「当該処理の意義および予測される結果」を提供しないといけないと規定している。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

④プロファイリングに基づく自動意思決定によって個人に不利な決定がなされた場合（雇用、ローン・保険など）、どのように対処すればよいか。

- 近年、AIを用いて新卒採用エントリーシートの判定を自動化したり、ローンやリース等の与信審査を自動化するといった事例が増えてきている。これらのケースで、AIの自動意思決定によって不利な結果を被ってしまった個人は、自動決定におとなしく従うしかないのだろうか。
- GDPRでは、本人に法的効果や重大な影響をもたらすような自動意思決定は原則として禁止されており、例外的に許される場合でも、「人を介在させる権利」、「自分の見解を表明する権利」、「決定に異議を唱える権利」を保障することが求められている。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

⑤プロファイリングによって個人の将来的なリスク（疾病、犯罪等）を予測する場合、予測データの「正確性」をどう捉えるか。

（1）プロファイリングによる予測データの正確性

- プロファイリングによって個人の疾病リスクなど将来的なリスクを予測するケースがあるが、予測は（単なる事実の推測と異なり）不正確な場合がありうる。入力データ（事実データ）が不正確な場合は本人の請求により訂正することができるが、出力データ（予測データ）が正確であるか否かはどのように判断したらよいか。また、本人は予測データを訂正することはできるのか。
- 例えば、ある機械学習アルゴリズムを用いて、個人Aの様々な健康データから「Aは心臓病に罹患するリスクが高い」ということが予測された。しかし、個人Aは5年経っても10年経っても心臓病に罹患しなかった。この場合、「心臓病に罹患するリスクが高い」という個人Aのプロファイル（予測データ）は正しいものなのか。また、この予測データのために高額の保険料を支払っていたり、保険申込みを拒否されていたりした場合、個人Aはデータの訂正を求めることはできるのか。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

⑤プロファイリングによって個人の将来的なリスク（疾病、犯罪等）を予測する場合、予測データの「正確性」をどう捉えるか。（続き）

（2）無意味な相関関係

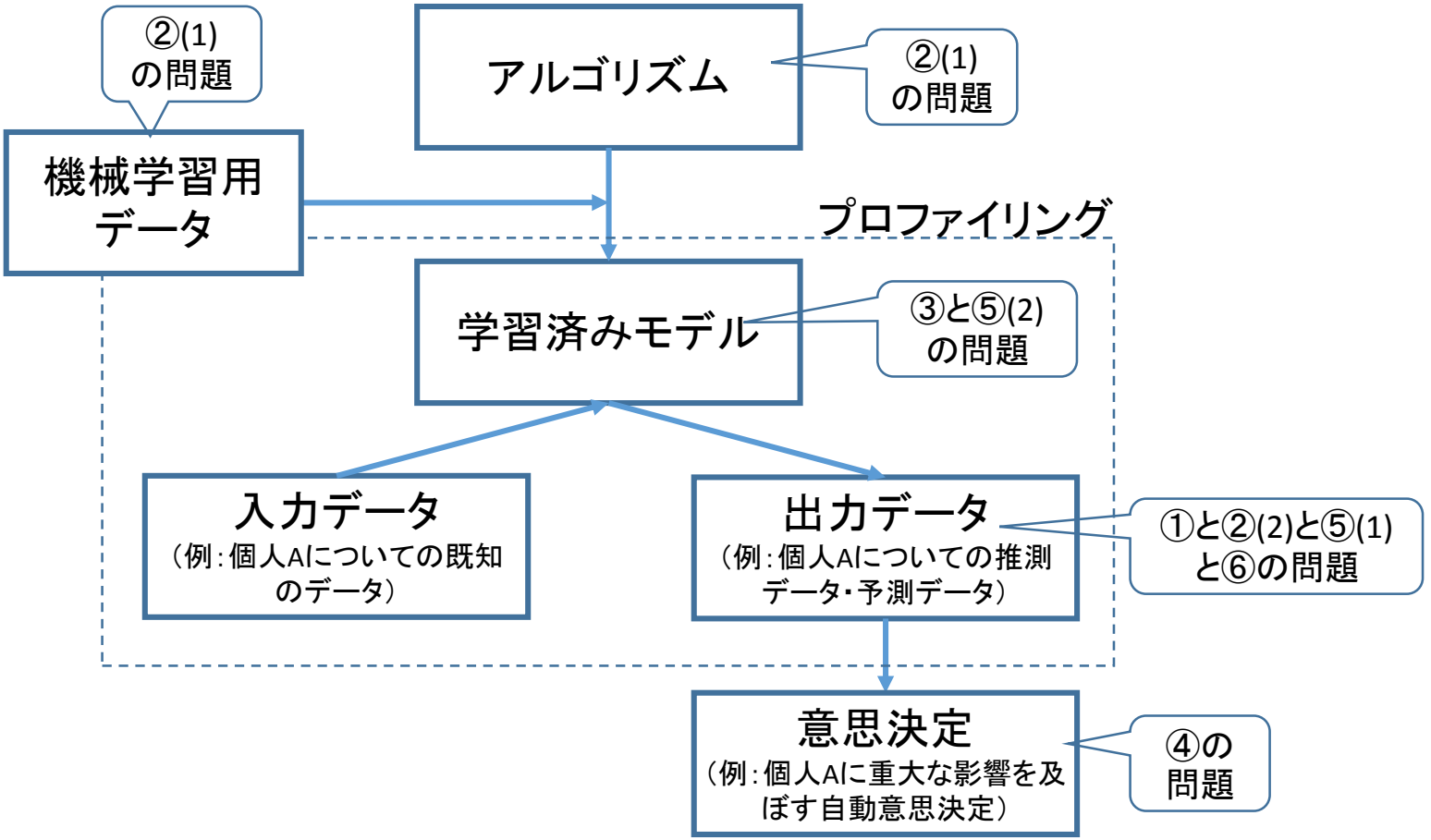
- ビッグデータ分析は相関関係の検出に非常に有効であるのに対し、どの相関関係が意味あるものであるかは説明しない。ある研究では「スイミングプールで溺れた人の毎年の数と、ニコラス・ケイジの出演作品の毎年の数」に相関関係があることが見出されたが、この相関が意味あるものとは思えない。
- しかし、仮に「ソーシャルメディアでの乱暴な表現方法と、自動車事故を起こすリスク」の間に相関関係が見出されたとした場合、それが意味のある相関関係か否かの判断（そして、その相関関係に従った意思決定を行うべきか否かの判断）は難しい。

⑥プロファイリングでパーソナライズされた環境によって本人の自律性（自律的な意思決定）が歪められることに、どのように対処すればよいか。

- 本人に重大な影響をもたらさないとしても、プロファイリングは本人の自律性を侵害する（自律的な意思決定を妨げる）可能性がある。例えば、ターゲティング広告やパーソナライズされたウェブコンテンツは、いわゆる「フィルターバブル」を生み出す恐れがある。管理者や本人は、このように自律性が侵害されることをどのように回避すればよいか。

プロファイリング・自動意思決定とプライバシー等に関する主な論点

- プロファイリング・自動意思決定とプライバシー等の問題



関連事例

①プロファイリングの結果、既知の個人情報から要配慮個人情報など本人の望まない個人情報が推測され、個人に紐付けられてしまったらどうするか。

○米国のスーパーマーケットTarget社

- 「無香料性ローション・特定サプリメント・大きめのバック」の購買履歴等から妊娠している可能性が高いと推測した顧客に、ベビー服等のオファーを提供した。ティーンエージャーの娘に対して、一緒に住む彼女の父親が妊娠の事実を知る前に、妊娠に関連する広告が送られてきてしまった。（出典：山本龍彦「ビッグデータ社会とプロファイリング」（『論究ジュリスト』2016年夏号所収））

○Facebookの「いいね！」を用いた研究

- ある研究では、Facebookの「いいね！」の履歴と他の情報を組合せることで、男性利用者の性的指向の88%、利用者の民族的素性の95%、利用者がキリスト教徒かイスラム教徒かの82%を正確に予測することができた。また民主党支持か共和党支持かの85%、アルコール・ドラッグ・煙草を使用しているかの65～75%を正確に推測することができた。（出典：Federal Trade Commission（FTC）「Big Data: A Tool for Inclusion or Exclusion?」（2016年1月））

○顔写真からセンシティブ情報を推測するプログラム

- 顔写真の特徴から、本人の信念や政治的見解、健康状態、性的指向まで推測するプログラムが存在するという。（出典：Chris Middleton「The Bias Virus」）

関連事例

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（1）AIアルゴリズムや機械学習用データにおける隠れたバイアス

〇米国ウィスコンシン州の再犯リスク予測評価システムCOMPAS

- 裁判官の量刑判断において、アルゴリズムを用いて犯罪者の再犯リスクを予測評価。COMPASが生成した再犯リスクスコアは、量刑判決、保釈、仮釈放の判断で利用される。ProPublicaの分析では、黒人の再犯リスクを白人の2倍に見積もっていた。2016年に同州最高裁判所がCOMPASを利用することについての合憲性判断を実施した。（出典：山本龍彦「ロボット・AIは人間の尊厳を奪うか？」（『ロボット・AIと法』（有斐閣、2018年4月）所収）、およびChris Middleton「The Bias Virus」）

〇スポーツジムのセキュリティシステム

- 英国の報告書では、ジムの自動セキュリティシステムが、女性の博士のタイトル「Dr」から当人を「男性」とプロファイリングしたため、ジムの更衣室からロックアウトされた事例が挙げられている。（出典：Information Commissioner's Office（ICO）「Big data, artificial intelligence, machine learning and data protection version 2.2」（2017年9月））

〇ある集団がビッグデータ分析のデータセットに過少に代表（反映）されていること

- ビッグデータ分析が依拠するデータセットから、一定の集団（population）に関する情報が欠落している場合がある。例えば、自分の個人情報を明かすことに注意深い個人、フォーマルな経済への参加が少ない個人、デジタルデバイドや「データ砂漠」によって技術へのアクセスが阻害されたり頻度が少ない個人、プロフィットの少ない顧客層として行動を単に観察されなかった個人など。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

関連事例

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（1）AIアルゴリズムや機械学習用データにおける隠れたバイアス（続き）

〇ボストン市の市民参加アプリ

- ボストン市はStreet Bumpというアプリを開発し、GPS機能を使って市民が道路状態・穴ぼこといった情報をレポートできるようにした。しかしアプリのリリース後、低収入の個人のスマートフォン所持率が低いため、アプリで得られるデータが全ての道路状態を代表（反映）している訳ではないことが判明。市政がこのようなバイアスのあるデータに依存し続けると、高収入世帯の地域に道路サービスが偏る恐れがある。ボストン市は、このアプリをまず市職員に利用させ、一般市民からのデータは補助的に使うことで対処した。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

〇ハリケーンSandy

- 2012年の10月27日から11月1日の間にハリケーンSandyに関するツィートが2000万件以上あった。このうち最大数のツィートはマンハッタンから発せられたものであり、あたかもマンハッタンが災害の中心地であるかのような幻想を生み出した。最も深刻な被害のあった場所から発せられたメッセージは非常に少なかった。なぜならこれらの地域ではスマートフォン所有率やツィッター利用率が低かったためである。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

関連事例

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（1）AIアルゴリズムや機械学習用データにおける隠れたバイアス（続き）

〇Web検索における検索結果

- ロイターやGoogleで黒人を示す名前（Jermaine等）を含むWeb検索を実行すると、白人を示す名前を含むWeb検索よりも、逮捕記録に関する広告表示が多いという米国の研究結果がある。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

〇顔認識技術における学習用データのバイアス

- 米国のMITでの研究は、広範に利用されている顔認識アルゴリズムは、白人の顔画像を主体として「トレーニング」されているために、バイアスがかかっていることを見出した。実験をした顔認識システムは白人男性の性別を99%の割合で正確に識別したが、黒人女性については35%であった 。英国ではビッグブラザーウォッチ（BBW）が最近（2018年）、警察に対する調査を報告した。その報告書では、ロンドン警視庁は自動顔照合において2%未満の正確性しかなく、102件の誤照合に対してたった2件の正しい照合しかなかったとしている。ロンドン警視庁はこの顔認識技術を用いて1人も逮捕者を出していない。（出典：英国庶民院科学技術委員会（Science and Technology Committee）「Biometrics strategy and forensic services」（2018年5月））

関連事例

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（1）AIアルゴリズムや機械学習用データにおける隠れたバイアス（続き）

〇マイクロソフトの対話エージェントTAY

- 米国マイクロソフトが開発した対話エージェントTAYは、差別的発言を繰り返すようになり、サービス開始後わずか2日で公開が中止された（2016年）。TAYにはユーザの発言を学習する機能があるが、悪質なユーザが意図的に悪い言葉を反復学習させてしまった。（出典：鳥海不二夫『強いAI・弱いAI』（丸善、2017年10月））

〇美人コンテストでのAI審査員

- 2016年9月の美人コンテストでAIプログラムが審査員になったとき、それはほとんどの黒人候補者を落選させた。なぜなら、「美人」を識別するための学習用データが十分な黒人女性を含んでいなかったためである。AIを学習させる人間が、無意識的にデータを偏らせていた。（出典：Chris Middleton「The Bias Virus」）

〇Google画像検索

- 2015年の時点で、「白人のティーンエージャー」としてGoogle画像検索すると、若い人々の写真の画像ライブラリーショットが検索結果として出るが、「黒人のティーンエージャー」としてGoogle画像検索すると、不均衡に高い割合で犯罪者・容疑者顔写真が出てくる。（出典：Chris Middleton「The Bias Virus」）

関連事例

②AIのアルゴリズムや機械学習用データにバイアスが含まれることにより、社会的差別が助長されたり新たに生み出されることをどう防ぐか。

（2）一定集団への「全体責任的」なプロファイル適用

〇クレジットカードの信用限度額

- ビッグデータ分析は、何らかの特徴を共有する他の消費者の行為に基づいた意思決定を導きうる。クレジットカード企業は、消費者自身の支払履歴ではなく、（同じ店で購入したことのある、支払履歴が芳しくない）他の消費者の分析に基づいてその人の信用限度額を低めることがある。実際、あるクレジットカード企業（CompuCredit Corp.）は、他の消費者の活動や支払履歴に基づいて、ある消費者が結婚カウンセリングやセラピー、タイヤ修理サービスにクレジットカードを使ったということで信用リスクが高いとレイティングしたことを開示しなかった容疑で、FTCへの陳述を行った。この種類の統計モデルを使うことはある個人の信用コストを下げるかもしれないが、信用力のある消費者が拒否されたり（統計モデルを使わなければ課されないような）金利を課されてしまう恐れもある。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

〇オンライン求職時の使用ブラウザ

- ビッグデータ分析は、企業がある集団（population）の人々を特定の機会から排斥することを正当化する口実を与えるかもしれない。例えば、あるビッグデータ分析の研究では「オンラインでの求職において、コンピュータに初期搭載されていないブラウザ（FirefoxやChromeなど自らインストールしなければならないブラウザ）を使って申請してきた人々は、よいパフォーマンスを示し、転職することも少ない」ことが示された。ある雇用主がこの相関関係を使って、特定のブラウザを使う人の採用を差し控えるとしたら、仕事と関係ない理由で質の高い求職者を排斥してしまうことになるかもしれない。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

関連事例

③アルゴリズムの不透明性（ブラックボックス化）に対する説明責任をどう考えるか。

○病気の予測システム

- 米国のマウントサイナイ病院は2015年、「ディープ・ペイシェント」という病気の予測システムを開発した。ディープ・ペイシェントは同病院が管理する7万6000人余りの電子カルテを読み込むと、そこに記されている身長・体重・血液・尿検査の結果などを分析し、これらの患者がいつ頃、どのような病気を発症したかを言い当てた。それらの病気は各種の癌や糖尿病、統合失調症など78種類に及ぶ。しかし、ディープ・ペイシェントは、ある患者がどんな病気を発症するかを予測できても、その根拠となる理由を教えてくれない。内部の思考回路が人間に見えないブラックボックスとなっている。（出典：小林雅一『AIが人間を殺す日』（集英社新書、2017年7月））

○ビッグデータにおけるプライバシーポリシー

- ビッグデータの分析手法は、個人が理解できる言葉で説明するのが難しすぎる。個人はそもそも長ったらしいポリシーを読みたがらない。IoTデバイスからデータを取得する場合、プライバシーポリシーの提示が難しい。ビッグデータ分析はしばしばデータの二次利用を伴うため、初めの時点で全ての利用目的を予見できないなど。（出典：ICO「Big data, artificial intelligence, machine learning and data protection version 2.2」）

○米国ウィスコンシン州の再犯リスク予測評価システムCOMPAS（再掲）

- ある黒人の被告人が、COMPASに基づき6年間の懲役および5年間の拡大保護観察を言い渡されたことについて、正確性が担保されず、検証可能性もない同システムで憲法が保障する適正手続の権利が侵害されたと提訴。取引の秘密などを理由にアルゴリズムがブラックボックス化されているため、「理由」が説明されないこと、かかる評価は同様の属性をもつ者が再犯を行う一般的な可能性を予測するにすぎないこと、黒人の再犯リスクを白人の2倍に見積もっていること等が問題とされた。（出典：山本龍彦「ロボット・AIは人間の尊厳を奪うか？」（『ロボット・AIと法』（有斐閣、2018年4月）所収））

関連事例

④プロファイリングに基づく自動意思決定によって個人に不利な決定がなされた場合（雇用、ローン・保険など）、どのように対処すればよいか。

○検索のオートコンプリート機能

- Google検索のオートコンプリート機能はドイツの裁判所で2回訴訟された。1つの訴訟では、Googleが「サイエントロジー（新宗教）」と「詐欺」をあるビジネスマンと結び付けることが訴えられた。もう1つの訴訟では、ドイツの前大統領の夫人からGoogleのオートコンプリート機能でエスコートサービスがサジェストされるとして訴えられた。（出典：ICO「Big data, artificial intelligence, machine learning and data protection version 2.2」）

○新卒採用のエントリーシート選考

- A社は2017年5月から新卒採用のエントリーシート選考にIBMのWatsonを活用。過去のエントリーシート選考のデータを学習したWatsonが、受験者のエントリーシートの合否判定をしているという。

○住宅ローンの仮審査

- B社は2018年5月から住宅ローンの仮審査において、独自に開発したAIを活用した自動審査の運用を開始。蓄積されたデータを定期的にAIに学習させることにより、審査担当者の判断に近い精度の自動審査を実現。また、B社では、これまで人手に頼っていた与信判断をAI が自動で行うことにより、業務の効率化と生産性の向上を実現。

○リースの与信審査

- C社は2018年１月、少額のリース案件をAI審査に切り替え、効率化できた分の人手を高額の審査などに振り向ける。

関連事例

⑤プロファイリングによって個人の将来的なリスク（疾病、犯罪等）を予測する場合、予測データの「正確性」をどう捉えるか。

（1）プロファイリングによる予測データの正確性

〇心臓病の罹患リスクの予測（架空ケース）

- 地域診療所のコンピュータシステムが或る個人を、「最も心臓病に罹りやすい」グループに分類した。この「プロファイル」は、たとえ当人が決して心臓病に罹らなかったとしても、必ずしも不正確なものではない。このプロファイルは単に当人が「最も心臓病に罹りやすい」と言っているだけである。これは、統計的な事柄として、事実上、正しいかもしれない。それにもかかわらず、データ主体は、処理の目的に鑑みて、補足的なステートメントを提供する権利を有する。上記のシナリオでは、制限された能力しかない地域診療所におけるコンピュータシステムよりも、追加的なデータを要素とすることができ、詳細な検査を実行できる、より先進的な医療コンピュータシステム（および統計モデル）に基づいたステートメントを提供することができる。（出典：EU指令第29条作業部会「Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)」）

関連事例

⑤プロファイリングによって個人の将来的なリスク（疾病、犯罪等）を予測する場合、予測データの「正確性」をどう捉えるか。

（1）プロファイリングによる予測データの正確性（続き）

〇シカゴ警察の犯罪予測システム

- シカゴ警察は、様々な経験的データ（例えば「殺人の犠牲者を知る者が自らも殺人に巻き込まれる可能性は通常の9倍である」といったデータ）から、暴力犯の実行者と犠牲者を予測するアルゴリズムを構築し、これを適用して潜在的な暴力犯実行者等をリスト化している。（出典：山本龍彦「ビッグデータ社会とプロファイリング」（『論究ジュリスト』2016年夏号所収））

〇遺伝子による人の将来の予測（架空ケース）

- 映画「Gattaca」（1997年）では、遺伝子操作により、優れた知能と体力と外見を持った「適正者」がエリート階層として社会で優遇され、自然妊娠で生まれた「不適正者」たちは遺伝子検査により出生時から知力や運動能力、職業適性、病気に罹るリスク、寿命等を算出され、入学・就職等を厳しく制限される。

関連事例

⑤プロファイリングによって個人の将来的なリスク（疾病、犯罪等）を予測する場合、予測データの「正確性」をどう捉えるか。

（2）無意味な相関関係

〇Googleのインフルエンザ・トレンド

- Googleの検索語に基づき、インフルエンザの流行を予測する機械学習アルゴリズム。米国でのインフルエンザの流行を予測するために、Googleチームは、特定地域でインフルエンザが発生していることを示す5000万の検索語を分析。当初は、このアルゴリズムはインフルエンザがどこで流行するかを正確に予測するように思われたが、その後非常に不正確な予測を行うようになった。これは、アルゴリズムがある変数を考慮に入れていなかったため。例えば、アルゴリズムは、インフルエンザの発生に関するストーリー（それが地球の裏側で起こったものだとしても）がローカルニュースで流れた際に、インフルエンザに関連した語が検索されるということを考慮に入れていなかったのかもしれない。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

〇ニコラス・ケイジの出演作品とスイミングプールで溺れた人の数

- Tyler Vigen氏の研究では、「スイミングプールで溺れた人の毎年の数と、ニコラス・ケイジの出演作品の毎年の数」には相関関係があるという。同氏はこの相関を、「偽の（うわべだけの）相関関係」（spurious correlation）と呼んでいる。他にも、「米国での毎年の日本車の販売数と、自動車事故での自殺者数」に相関関係があるという。（出典：山本龍彦「ロボット・AIは人間の尊厳を奪うか？」、Tylervigen.com「Spurious Correlations」）

〇ソーシャルメディアと自動車事故のリスク

- AIで自動車保険の料金を決める場合、ソーシャルメディアのデータを用い、書き込み内容ではなく、書き方や表現方法だけに注目したケースがある。すなわち、「！」の数やエキサイトした書き方のみで、自動車事故のリスクが高いと決定していた。このアルゴリズムは実際にはパイロットのみで、実際には使われなかった。（出典：Information Commissioner's Office（ICO）へのヒアリング）

関連事例

⑥プロファイリングでパーソナライズされた環境によって本人の自律性（自律的な意思決定）が歪められることに、どのように対処すればよいか。

〇悪徳商法の提供（架空ケース）

- ビッグデータ分析は、脆弱な消費者が詐欺の標的になるのを助長しうる。悪徳企業はビッグデータを利用して、最も脆弱な見込み客に対して、ミスリーディングなオファーや悪徳商法を提供することができる。公的な報告書によれば、悪徳な企業は販売くじのオファーに応えた人々（したがって誘惑に最も乗りやすい人々）のリストや、アルツハイマー症などの病気に罹っている「苦しんでいる高齢者」のリストを取得することができる。ビッグデータ分析によって、企業はより容易かつ正確にそのような脆弱な見込み客を特定することができる。（出典：FTC「Big Data: A Tool for Inclusion or Exclusion?」）

〇個別化広告の逆影響

- オハイオ州立大学の研究チームは、個人の趣味嗜好をプロファイリングして送られる個別化広告が個人の「自意識」自体に重大な影響を与えることを例証。被験者となった大学生は自分のオンライン上の行動の結果として送られてくる個別化広告（例えば環境保護的な広告）を、それが本当に自分の性向とマッチしているか否かにかかわらず、「自己の反映」として（例えば自己を環境保護に熱心な人間として）認識する傾向がある。AIの予測評価は本来、共通した属性をもつ集団の一般的傾向を示しているにすぎないが、そうなるとAI社会では我々の個人的なアイデンティティがある特定の集団の一般的傾向によって逆規定されるようになる可能性がある。（出典：山本龍彦「ロボット・AIは人間の尊厳を奪うか？」（『ロボット・AIと法』（有斐閣、2018年4月）所収））

関連事例

⑥プロファイリングでパーソナライズされた環境によって本人の自律性（自律的な意思決定）が歪められることに、どのように対処すればよいか。

〇ケンブリッジ・アナリティカ社の政治広告

- 2016年の米国大統領選挙でトランプ陣営は、ビッグデータにケンブリッジ・アナリティカ社が開発した有権者の心理分析を加え、小グループごとに向けて開発した「個別広告（マイクロターゲット広告）」を、特定の地域でテレビ、電子メールやソーシャルメディアを通じて投入した。マイクロターゲット広告とは、有権者のパーソナリティーに応じてニュアンスや広告のタイプを変える、個別化されたプロパガンダのようなものである。例えば、政府による銃規制に不安をもっている40代の父親には「父から息子へ」と題して夕暮れの野原で射撃訓練をする親子の映像を用い、トランプ候補が当選すれば必ず銃所有の権利を守ってくれるという広告を送る。また、空き巣に対する不安を抱いている30代のシングル女性には、窓が割られた住宅の写真を添え、トランプ候補であれば、治安改善に最善の努力を尽くすという広告を送る。たとえ夫婦であっても、心理分析の結果次第で別のタイプの広告をそれぞれのソーシャルメディアへ送るという。（出典：福田直子『デジタル・ポピュリズム 操作される世論と民主主義』（集英社新書、2018年5月））

２．欧州一般データ保護規則（GDPR）におけるプロファイリング規制

GDPRにおけるプロファイリング（第22条等）

- GDPRではプロファイリング（※）を以下の3つの場面で規制
  - ※プロファイリングとは：「個人を一定のカテゴリーに分類したり、個人の遂行能力・興味・行動等について分析や予測をするために、個人に関する情報を集めて、その特徴や行動のパターンを評価すること」（第29条作業部会ガイドライン（WP251））

プロファイリングの種類	説明（GDPRでの位置付け）	例（第29条作業部会ガイドライン（WP251）記載のもの）
①プロファイリング一般	第4条4項で定義。個人データ処理の1つとして（他の個人データ処理と同様に）様々な義務。	・データブローカーが様々な情報源から個人データを取得し、整理して、個人に関するプロファイルを作成し、セグメント分け（し、顧客企業に販売）する場合
②プロファイリングなどに基づく自動意思決定（第22条）	プロファイリングなどに基づく完全に自動化された意思決定であって、本人に法的効果または同様の重大な影響をもたらすもの。第22条で原則として禁止。（契約履行に必要な場合や、本人の明示的同意がある場合は可能だが、人を介在させる権利、自分の見解を表明する権利、決定に異議を唱える権利を保障しないといけない。）	・個人の運転習慣が長期に渡ってモニターされ、「常習的なスピード違反か」「直近に他の交通違反を起こしていないか」等に基づいて（自動的に）罰金額が決められる場合 ・ローンの審査がアルゴリズムを用いて行われ、担当者による評価を経ずに、審査結果が自動的に個人に通知される場合 ・オンラインでのクレジットカード申請の自動的な拒否 ・人間が介在しない電子リクルーティング
③ダイレクトマーケティング目的でのデータ処理（プロファイリングなど）（第21条）	ダイレクトマーケティング目的での個人データ処理（プロファイリングなど）に対して、本人はいつでも異議を唱える権利を持つ（第21条2項）。（公共の利益や正当な利益の目的での個人データ処理（プロファイリングなど）に対して異議を唱える権利もある（第21条1項）。）	・利用者の近くのレストランをレコメンドする携帯アプリが、取得したデータから利用者の食事の好みや生活習慣などをプロファイリングし、携帯電話に広告を送る場合。

（出典：国際社会経済研究所）

GDPRの条文（第22条）

第22条 プロファイリングを含む自動化された個人意思決定

- 1. データ主体は、当該データ主体に関する法的効果をもたらすか又は当該データ主体に同様の重大な影響をもたらす、プロファイリングなどの自動化された処理のみに基づいた決定に服しない権利を持つ。
- 2. 第1項は次に掲げるいずれかの決定には適用されない。
  - (a) データ主体とデータ管理者間の契約締結、又は履行に必要な決定。
  - (b) データ主体の権利及び自由並びに正当な利益を保護するための適切な措置が定められた管理者が従うEU 法又は加盟国の国内法によって認められた決定。
  - (c) データ主体の明示的な同意に基づく決定。
- 3. 第2項(a)号及び(c)号で定める状況に関して、データ管理者は、データ主体の権利及び自由並びに正当な利益を保護するための適切な措置を実施し、少なくとも管理者側で人を介在させる権利、当該データ主体の見解を表明する権利、及び決定に異議を唱える権利を実施するものとする。
- 4. 第2項で定める決定は、第9条第2項(a)号又は(g)号が適用されず、データ主体の権利及び自由並びに正当な利益を保護するための適切な措置が機能していないならば、第9条第1項で定める特別な種類の個人データに基づいてはならない。

出典：JIPDECのGDPR仮日本語訳（訳語を一部変更、下線を記入）

GDPRの条文（第13条）

第13 条 データ主体から個人データを収集する場合に提供される情報

1. データ主体に係る個人データがデータ主体から収集される場合、管理者は、個人データを取得する際、データ主体に次に掲げるすべての情報を提供するものとする。

(a) 管理者の身元及び詳細な連絡先、該当する場合、管理者の代理人。

(b) 該当する場合、データ保護オフィサーの詳細な連絡先。

(c) 意図された個人データの取扱い目的、及び取扱いの法的根拠。

(d) 取扱いが第6 条第1 項(f)号に基づく場合、管理者又は第三者によって求められる正当な利益。

(e) もしあるならば、取得者又は個人データの取得者の種類。

(f) 該当する場合、管理者が個人データの移転を意図する第三国若しくは国際組織及び欧州委員会による十分性決定の存在若しくは不存在に関する事実、又は第46 条若しくは第47 条、若しくは第49 条第1 項後段で定める移転状況において、適切若しくは適正な保護措置及び個人データの複製を取得する方法若しくは個人データが利用可能になる条件に関する情報。

2. 第1 項で定める情報に加え、管理者は、個人データを取得する際、公正で透明性のある取扱いを保障するために、データ主体に次に掲げる必要な追加的情報を提供するものとする。

(a) 個人データが保存される期間、もし不可能であるならば、当該期間を決定するのに用いられる基準。

(b) 管理者に対し個人データへのアクセス、訂正又は消去、データ主体についての取扱いの制限を要求する権利、又は当該取扱いに不服を申し立てる権利とともに、データポータビリティーの権利の存在。

(c) 取扱いが第6 条第1 項(a)号又は第9 条第2 項(a)号に基づく場合、撤回前の同意に基づく適法な取扱いに影響を与えることなしに、いつでも同意を撤回する権利の存在。

(d) 監督機関に不服を申し立てる権利。

(e) 個人データの提供が、法令又は契約上の要件、又は契約を締結するのに必要な要件であるか否か、及びデータ主体に個人データの提供の義務があるか否か、並びに当該データ提供の不履行により起こり得る結果。

(f) プロファイリングを含め、第22 条第1 項及び第4 項で定める自動化された意思決定の存在、少なくともそのような状況において、関連する論理について意味ある情報、データ主体に関する当該処理の意義及び予測される結果。（以下、第3項、第4項は略）

出典：JIPDECのGDPR仮日本語訳（訳語を一部変更、下線を記入）

自動意思決定とプロファイリングに関するガイドライン：定義

- 以下、EU指令第29条作業部会「自動化された意思決定とプロファイリングに関するガイドライン」（WP251rev.01）（2018年2月）の概要である。

A. プロファイリング（Profiling）

第4条(4)：

「プロファイリング」とは、自然人に関するある一定の個人的な側面を評価するために、特に、自然人の業務実績、経済状況、健康、個人的嗜好、興味、信頼、行動、所在又は移動に関連する側面の分析又は予測をするためになされる、個人データの利用から成る個人データのあらゆる形態の自動的な処理をいう。

- 上記定義により、プロファイリングは以下の3つの要素から成る。
  - 自動的な形態の処理（を伴ったもの）。（人間の関与を必ずしも排除しない。）
  - 個人データに対して実施される処理。
  - 自然人に関する個人的な側面を評価する処理。

※ 欧州評議会の2010年の勧告では、プロファイリングを以下の3段階に分けて記述している。
  - ①データ収集。
  - ②相関関係を特定するための自動的な分析。
  - ③相関関係を個人に適用し、現在または将来の行動の特徴を特定すること。

- プロファイリングは、「個人を一定のカテゴリーに分類したり、個人の遂行能力・興味・行動等について分析や予測をするために、個人に関する情報を集めて、その特徴や行動のパターンを評価すること」を意味すると概括されている。

自動意思決定とプロファイリングに関するガイドライン：定義

B. 自動意思決定（Automated decision-making）

- 「自動意思決定」は「プロファイリング」と一部重なる場合もあるが、異なる概念である。
  - 例： スピードカメラのみに基づいて科される罰金は自動意思決定であるが、プロファイリングは伴わない。
  - 例： 個人の運転習慣が長期に渡ってモニターされ、「常習的なスピード違反か」「直近に他の交通違反を起こしていないか」といった要素に基づいて罰金額が決められる場合、プロファイリングに基づく意思決定となる。
- 自動意思決定は以下のデータに基づいてなされる。
  - 関係する個人から直接的に提供されるデータ（質問フォームへの回答など）
  - 個人を観察することで得られるデータ（アプリを通じて収集される位置データなど）
  - 既存のプロファイル情報から推測されたデータ（信用スコアなど）

C. GDPRにおける「プロファイリング」の扱い

- 「プロファイリング」という概念は、以下の3つの使われ方がなされうる。
  - (i)プロファイリング一般
  - (ii)プロファイリングに基づく意思決定（人間が関与するもの）
  - (iii)プロファイリングなど（に基づく）、自動意思決定（第22条）
- (iii)には、GDPR第22条が適用される。（同ガイドラインの中心テーマ）
- (i)(ii)には、GDPR全般（一般原則）が適用される。

自動意思決定とプロファイリング： 概要図

当該個人データ処理は、自動化された処理のみ（プロファイリングなど）に基づく意思決定か？

No → GDPR全般が適用される（通常の個人データ処理）

Yes ↓

当該意思決定は、本人に法的効果または同様の重大な影響をもたらすか？

No → GDPR全般が適用される（通常の個人データ処理）

Yes ↓

当該意思決定は以下のいずれかに該当するか？

(a)本人と管理者間の契約締結、又は履行に必要な決定。
(b)本人の権利及び自由並びに正当な利益を保護するための適切な措置が定められた管理者が従うEU 法又は加盟国の国内法によって認められた決定。
(c)本人の明示的な同意に基づく決定。

No → 原則として禁止

Yes ↓

- 本人への情報提供において「自動意思決定の存在」「関連するロジックについて意味のある情報」「当該処理の意義および予測される結果」を提供する義務（第13条、第14条）
- 本人の権利及び自由並びに正当な利益を保護するための適切な措置を実施する義務（上記(a)と(c)の場合）（第22条3項）。少なくとも
  - 管理者側で人を介在させる権利
  - 本人の見解を表明する権利
  - 決定に異議を唱える権利
- データ保護影響評価（DPIA）を実施する義務（第35条3項） 等

（出典：国際社会経済研究所）

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ 第22条は以下を意味する。

- ①「データ主体に法的効果または同様の重大な影響をもたらす、プロファイリングなどの完全に自動意思決定」を原則的に禁止する。（第22条1項）
- ② この原則にはいくつかの例外がある。（第22条2項）
- ③ 例外の適用に当たっては、データ主体の権利及び自由並びに正当な利益を保護するための適切な措置を実施するものとする。（第22条3項）

○ 第22条2項で挙げられた3つの例外

(a) データ主体とデータ管理者間の契約締結、又は履行に必要な決定。

- 管理者は、それが目的を達成する最も適切な方法だという理由から、自動意思決定を使用したいと思うかもしれない。人間が定常的に関与することは、処理されるデータ量の観点から非現実的であるか不可能であるかもしれない。
- このような自動意思決定が必要と示すためには、よりプライバシー侵害的でない手段を採用しうるかどうかを考慮しなければならない。もし他の手段によって同じ目的に到達できるのであれば、このような自動意思決定は契約履行等に「必要」とはみなされない。

(b) データ主体の権利及び自由並びに正当な利益を保護するための適切な措置が定められた管理者が従うEU 法又は加盟国の国内法によって認められた決定。

- 前文71で挙げられている例は、以下のために自動意思決定を使用することがEU法または加盟国法で定められている場合。
  - 詐欺の監視や防止
  - 脱税の監視や防止
  - サービスのセキュリティの保証

(c) データ主体の明示的な同意に基づく決定。

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ 第22条1項の「自動化された処理のみに基づいた」の意味：

- これは、意思決定プロセスに人間の関与がないことを意味する。

○ 第22条1項の「法的効果」または「同様の重大な影響」の意味：

「法的効果」

- 自動意思決定がデータ主体に「法的効果」をもたらす例として、以下。
  - 契約を解除されること。
  - 法令で定められた社会保障給付金（児童手当や住宅手当等）の受給資格を得たり、拒否されること。
  - 入国を拒否されたり、市民権を拒否されること。

「同様の重大な影響」

- 前文71で挙げられている例は、「オンラインでのクレジットカード申込みの自動的な拒否」および「人間が介在しない電子リクルーティング」。
- この条件を満たすには、当該処理の効果が些細なものであってはならず、注目に値するほど十分に大きいか重要なものでなければならない。換言すれば、当該意思決定は以下の可能性がなければならない。
  - 個人の環境や行動、選択に重要な影響を与えること。
  - データ主体に長期的または永続的な影響を与えること。
  - 極端な場合、個人の排斥や差別につながること。
- 十分に「重大」な影響であるか否かの閾値を正確に示すことは難しいが、以下の決定はこのカテゴリーに入りうる。
  - 信用度など、当人の金融的環境に影響を与える決定
  - 当人の医療サービスへのアクセスに影響を与える決定
  - 当人の雇用機会を否定する決定、または当人を重大な不利な状況に追いやる決定
  - 大学入学など、当人の教育へのアクセスに影響を与える決定

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ データ主体の権利： 情報提供を受ける権利（第13条、第14条）

- 管理者が第22条1項にいう自動意思決定を行う場合は、
  - データ主体にその旨を告知しなければならない。
  - 関連するロジック（logic involved）について意味のある情報を提供しなければならない。
  - 当該処理の意義（significance）および予測される結果について説明しなければならない。
- また、当該処理が第22条1項の狭い定義に当てはまらない場合でも、上記の情報を提供することはグッドプラクティスである。
- 特にプロファイリングに基づく意思決定である場合には（第22条1項の定義に当てはまるか否かに関わらず）、当該処理が「プロファイリングを目的としていること」および「生成したプロファイルに基づいて意思決定を行うことを目的としていること」の両方をデータ主体に明確に示さなければならない。

- 「当該処理の意義（significance）および予測される結果」：
  - 例： ある保険会社は、顧客の運転行動のモニタリングに基づいて、自動車保険の保険料を自動決定している。この自動意思決定の「意義」と「予測される結果」として、保険会社は「危険な運転は高い保険料支払いにつながる可能性がある」と説明し、架空ドライバー（急加速したり、ブレーキが遅いドライバーなど）の保険料を比較するアプリを提供。また、図面を使って、どのように運転習慣を改善すれば保険料が下がるかのヒントを説明。

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ データ主体の権利： 情報提供を受ける権利（第13条、第14条）

- 「関連するロジックについての意味のある情報」：
  - 機械学習の発展により、自動意思決定プロセスやプロファイリングがどのように行われているかを理解することは難しくなっている。
  - 管理者は、意思決定の背後にある理由付けや依拠するクライテリアについてデータ主体に伝えるシンプルな方法を見出すべきである。また、データ主体に提供される情報は意味あるものであるべきである。
  - 例： 管理者は信用スコアを個人のローン申請を評価したり拒否したりするために利用している。信用スコアは信用照会機関から提供されたものだったり、管理者が保持する情報に基づき直接的に計算されたものだったりしうる。これらの情報源に関わらず、管理者がこの信用スコアに依存しているならば、管理者はその旨と根拠について、データ主体に説明できねばならない。管理者は、このプロセスが公正で責任ある貸付決定を行うことに役立つことを説明する。管理者は決定において考慮される主な項目の詳細、当該情報の情報源、その関連性について提供する。これには、例えば以下が含まれる。

    ・データ主体によって申請フォーム上で提供された情報

    ・これまでの口座取引に関する情報。支払滞納に関する情報を含む

    ・詐欺記録や支払不能（破産）記録などの公的な記録情報

    管理者はまた、使用されている信用スコア手法の公平性・有効性・バイアスがかかっていないことを保証するためにそれが定期的にテストされていることもデータ主体に情報提供する。管理者は、GDPR第22条3項に則り、ローン申請を拒否されたデータ主体が決定の再考を要求できるように連絡先詳細を提供する。

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ 適切な措置（第22条3項）

- 第22条3項の「データ主体の権利及び自由並びに正当な利益を保護するための適切な措置」には、少なくとも、データ主体が人の介在を得るための方法、自分の見解を表明するための方法、当該意思決定に異議を唱えるための方法を含めるべき。

- 人の介在：
  - いかなるレビューも、意思決定を変更する適切な権限（authority）と能力（capability）を持った人間によって実施されなければならない。このレビュワーは、データ主体から提供される追加情報を含め、全ての関連するデータの徹底的な評価を実施するべきである。

- 透明性：
  - 前文71では、第22条3項に列挙された権利に加え、「データ主体に対する特別の情報提供、そのような評価の後に到達した決定について説明を受ける権利」を保障すべきとしている。これは、当該処理に関する透明性の必要性を強調するものである。本人は、意思決定がどのように、何を根拠になされたかを完全に理解した場合のみ、その意思決定に異議を述べたり自分の見解を表明することが可能になるだろう。

- バイアスやエラーへの対処：
  - 管理者は、処理するデータセットに対して、何らかのバイアスがないかをチェックするために、頻繁な評価を実施するべきである。そして、相関関係への過剰な依存を含め、偏見・先入観のある要素に対処する方法を開発するべきである。アルゴリズムを監査するシステムや、プロファイリング等の自動意思決定の正確性・関連性の定期的なレビューは、その他の有益な措置である。管理者は、特別な種類のデータに基づくエラーや不正確性、差別を防ぐための適切な手続きや措置を導入するべきである。これらの措置は、設計段階のみならず、このプロファイリングが個人に適用されるまで、継続的に用いられるべきである。そのようなテストの結果は、当該システムの設計にフィードバックされるべきである

自動意思決定とプロファイリングに関するガイドライン：第22条関連

○ グッドプラクティスの勧告： 第22条と前文71における適切な措置（safeguards）

- **システムの定期的な品質保証チェック**： 特別な種類の個人データに基づくか否かにかかわらず、個人が公正に取り扱われており、差別されていないことを確認する。
- **アルゴリズムの監査**： 機械学習によって使用され開発されたアルゴリズムをテストし、それらが意図されたとおりに実際に動いて(perform)おり、差別的、誤りのある、あるいは正当化できない結果を生み出していないことを証明する。
- **独立の第三者による監査**（プロファイリングに基づく意思決定が個人に高い影響を及ぼす場合）： 監査者にアルゴリズムや機械学習システムがどのように機能（work）するかに関する全ての必要な情報を提供する。
- **第三者のアルゴリズムに対する契約上の保証**： 監査とテストが実施され、アルゴリズムが合意された基準を遵守していることの契約上の保証を得る。
- **データ最小化のための特別な措置**： プロファイル自体、およびプロファイルを作成したり、プロファイルに適用する際に利用される個人データについて、明確な保持期間を組み込む。
- プロファイリングの文脈における**匿名化技術または仮名化技術の使用**
- データ主体が**自分の見解を表明したり、意思決定に異議を申し立てることを可能にする方法**
- 既定のケースにおける**人間の介在のためのメカニズム**： 例えば、データ主体に自動意思決定が伝えられる時点でのアピール（異議申立）プロセスへのリンクの提供。レビューのための合意されたタイムスケールや、質問に対するコンタクトポイントを伴うもの。

- 管理者は以下のようなオプションを検討することもできる。
  - 認証メカニズム
  - 機械学習に関連した監査プロセス向けの行動規範
  - 倫理レビュー委員会： プロファイリングが適用される分野への潜在的な害悪とベネフィットを評価する。

自動意思決定とプロファイリングに関するガイドライン：その他

○ 子どもと自動意思決定

- 前文71で、データ主体に法的効果または同様の重大な影響をもたらすプロファイリングなどに基づく完全に自動化された意思決定は、子どもに適用しないものとするとされている。条文本文（第22条）には子どもに関する言及はないものの、第29条作業部会は子どもに対するこのような自動意思決定を正当化するために第22条2項の例外を適用しないことを推奨している。

○ データ保護影響評価（DPIA）との関連

- 第35条3項(a)では、プロファイリングなどに基づく自動意思決定に対するDPIAの実施義務が規定されている。

○ 特別な種類の個人データ（第9条）との関連

- プロファイリングによって派生したデータや推測データに「特別な種類の個人データ」（要配慮個人情報に相当）が含まれる場合がありうる。この場合、「特別な種類の個人データ」として処理しなければならない。
  - 例：食品購買履歴から当人の健康状態を推測する場合。
  - 例：ある調査研究では、Facebookの「いいね！」の履歴と他の情報を組合せることで、男性利用者の性的指向の88%、利用者の民族的素性の95%、利用者がキリスト教徒かイスラム教徒化の82%を正確に予測することができた。

３．英国ICOヒアリング調査
（プロファイリング・自動意思決定）

英国ICO調査概要

- 概要
  - 2018年7月13日に英国のデータ保護監督機関である情報コミッショナーオフィス（ICO）を訪問し、Policy and Engagement Group、Technology GroupおよびInternational Groupの方々に、プロファイリングおよび自動意思決定に係る諸問題についてヒアリングを行った。質問表は先方に事前送付した。

- 面会者
  - Mr. Carl Wiper (Group Manager, Policy and Engagement)
  - Ms. Karen Harris (Senior Policy Officer, Policy and Engagement)
  - Mr. Nigel Houlden (Head of Technology)
  - Mr. Alain Kapper (Senior Policy Officer, International Group)
  - Ms. Fabiana Marinaro (Lead Policy Officer, International Group)

調査の背景

〇パーソナルデータの利活用に関する制度改正大綱

- 国内法ではプロファイリング等に関する明示的な規制はないが、「パーソナルデータの利活用に関する制度改正大綱」（2014年6月）において、「継続的な検討課題」として以下が挙げられている。
  - １ 新たな紛争処理体制の在り方
  - ２ いわゆるプロファイリング：

    多種多量な情報を、分野横断的に活用することによって生まれるイノベーションや、それによる新ビジネスの創出等が期待される中、プロファイリングの対象範囲、個人の権利利益の侵害を抑止するために必要な対応策等については、現状の被害実態、民間主導による自主的な取組の有効性及び諸外国の動向を勘案しつつ、継続して検討すべき課題とする。
  - ３ プライバシー影響評価(PIA)
  - ４ いわゆる名簿屋

〇改正個人情報保護法

- また、改正個人情報保護法の附則第12条第3項において、以下のように規定されている。改正法は2017年5月施行なので、「施行後三年」は2020年に当たる。
  - 「政府は、前項に定める事項のほか、この法律の施行後三年ごとに、個人情報の保護に関する国際的動向、情報通信技術の進展、それに伴う個人情報を活用した新たな産業の創出及び発展の状況等を勘案し、新個人情報保護法の施行の状況について検討を加え、必要があると認めるときは、その結果に基づいて所要の措置を講ずるものとする。」
- EUでは旧データ保護指令およびGDPRにおいてプロファイリングは規制対象となっており、「個人情報の保護に関する国際的動向」やEUからの十分性認定を踏まえると、わが国でもプロファイリングに対する規制が「見直し」の結果として、個人情報保護法に追加される可能性も想定される。

（１）プロファイリングの定義／範囲

- Q1.1：プロファイリングの語は、GDPRやICOのガイドではやや広い意味で用いられているように思える。例えば、自動車メーカーのサイトを訪れたことのある人に対し、その閲覧履歴に基づいて広告企業が他のサイトで自動車のターゲティング広告を表示したとする。この場合も「プロファイリング」と言うのか？
- Q1.2：それとも、「このメーカーの自動車に興味がある」といったプロファイルが作成された（そのようなラベル付けがされた）ときに初めて「プロファイリング」に該当するのか？

- A：質問1.1の事例は単純でベーシックなプロファイリングの例に当たる。GDPRの前文30で、どの識別子が個人プロファイルを作りうるかのリストが挙げられている（IPアドレス、クッキー、RFIDタグなどの識別子）。法的には、質問1.1と1.2の事例を分けて考えにくい。根本的には同じものである。1.1で自動意思決定が一瞬で行われたとしても、同じである。

- Q：GDPRのプロファイリングガイドラインでは、スピードカメラで罰金を自動決定することはプロファイリングではないとされている。

- A：確かに決定は自動的であるが、映像による証拠など単なる事実に基づいており、プレディクション（予想）をしていない。すなわち、その個人の趣味や行動などの分析を行っていない。スピード違反の事例でも、もし年齢・車種など他の情報と組み合わせ（て罰金を決める）のであれば、プロファイリングに当たる。

（２）プロファイリングと特別な種類のデータ（センシティブデータ）

- Q2.1：GDPRのプロファイリングガイドラインによれば、通常の個人データからプロファイリングによる推測によって、特別な種類のデータ（いわゆるセンシティブデータ）を作成してしまう場合がある。この場合、管理者は実務上、どのようにしてこれらの特別な種類のデータを合法的に取り扱えばよいのか？
- （プロファイリングに先立って）前もってデータ主体から（特別な種類のデータを取扱うことについて）明示的な同意を取得しておいてもよいのか？

- A：特別な種類のデータを推測できるのは事実である。管理者側に、推測した特別な種類のデータを必要とする目的がなければ、保持してはいけない。逆に利用したい目的があるならば、合法的な根拠（GDPR第6条）のいずれかがなければいけない。本人の明示的な同意は、合法的な根拠の1つの例にすぎない。他にも合法的な根拠がある。英国のデータ保護法2018では、合法性の根拠としてGDPRに加えて追加的な根拠を規定している。

（３）プロファイリングと本人の望まないデータ推測

- Q3.1：センシティブデータ以外でも、本人が管理者への提供を拒否した個人データ項目について、プロファイリングによる推測によって、管理者は作成（入手）することができるかもしれない。そのような、本人が期待しないプロファイリングから個人を守る手段はないのか？

- A：これは、データ処理の透明性に関わる問題である。推測した場合、データ主体にとって目に見えないのが問題だ。GDPRの透明性の要件を満たすことが重要である。
  - 既知の個人データから「推測」することをデータ主体に伝える。
  - 異議を唱える権利などを明示する。
- 例えば、英国の小売店が発行するメンバーカードで購買履歴を把握する場合、消費者はディスカウントや特別オファーがくるために購買履歴を使っていると考える。もし、メンバーカードで集めたデータが、ローン審査の拒否や保険料のアップに使われたとしたら、消費者は困る。こうした場合、データ主体が予期しない使われ方になる。（英国ではスーパーが保険を販売することもある。）

- Q：その前提として、企業が正直でないと駄目なのではないか？

- A：透明性の要件は単なるグッドプラクティスではなく、法律で義務付けられている。管理者が遵守しなければ、ICOの役務である。こっそりプロファイリングするような企業があれば、ICOが摘発することになる。
- 2018年7月11日に、ICOはFacebookとケンブリッジ・アナリティカ社に関するレポート を出した。これはまさに、企業がデータ主体に秘密でプロファイリングを行った例である。

（４）自動意思決定におけるデータ主体の権利（１/４）

- Q4.1：GDPR第22条2項の自動意思決定が本人に不利な結果をもたらした場合（人事採用など）、どこまで本人の異議申立ては許容されるのか？

- A：自動意思決定に対する異議申立は可能である。まず、自分の見解を表明する。もっと大事なのは、人の介在を要求できる権利である。これは考え直してもらう権利とも言える。データ主体が管理者に対して異議申立をしても、その管理者に対して権利を行使できないと思ったら、DPAに申し立てることができる。直接、裁判所に訴えることもできる。
- 第22条に自動意思決定に関する規定がある。全ての自動意思決定が規制されている訳ではなく、「完全に自動化」「本人に法的効果または同様の重大な影響をもたらす」という2つの条件がある。

（４）自動意思決定におけるデータ主体の権利（２/４）

- Q4.2：意思決定プロセスに何らかの誤りがあった場合（例えば、意思決定が本人に関する誤ったデータに基づいて行われた場合）のみ、訂正することができるのか？

- A：GDPR第22条が与えている権利は、自動意思決定に異議を申し立てる権利であって、当該決定を変更する権利ではない。GDPRで規制している目的としては、自動意思決定が増えているので、問題視しており、人の介在や異議申立（contest）の権利を与えている。
- 人が介在しても、コンピュータの自動意思決定と一緒の結論になる可能性がある。決定が同じだったとしても、データ保護とは関係がない。データ保護上で保障されているのは、決定に対して「おかしい」と言える権利だけである。
- 後からデータ主体が追加で自分のデータを提供することで、決定が変更される場合がありうる。典型的な例として、ローンの合否を決めるのにアルゴリズムで自動で決定するのが普通になっている。異議申立できる権利をローン会社が明示しなかったとしたら、データ保護法（GDPR）違反である。この場合は、ICOのマターになる。

（４）自動意思決定におけるデータ主体の権利（３/４）

- Q4.3：人間もまたバイアスを持ったり、誤りを行いうる。なぜ自動処理のみに基づく意思決定は原則的に「禁止」され、人間を介在させるとOKなのか？

- A：確かに人間も間違いうるので、アルゴリズムに任せた方が人間がやるよりも決定しやすい場合もあるだろう。しかし、GDPRが問題視しているのは、どんどん自動意思決定なされる場面が増えているということである。人間に介在してもらう権利を回復することで、データ主体が「おかしい」と思った際に、単に機械だけに判断されるのではなく、人間と対話をすることができる。
- 最近あるイベントに参加した。「Citizen Jury」（一般市民の陪審員）向けのイベント。英国の研究所が主催し、テーマは「自動意思決定とプロファイリング」である。一般市民のディスカッションがあった。「人間と対話できるべきだ」という意見がたくさん出た。

（４）自動意思決定におけるデータ主体の権利（４/４）

- Q4.4：GDPR第22条2項の自動意思決定で本人に不利な結果となり、人間の介在を要求したが、結果が覆らなかった場合、本人はさらに異議申立できるのか？

- A:データ主体にとって不利な決定だったとしても、それはGDPRの話ではない。GDPRでは異議申立（contest）の権利を与えることが重要であり、決定が変更される保証がある訳ではない。
- データ主体が、管理者がきちんと法律の要件に従っていないと思ったら、ICOに連絡することができる。ICOは調査して、是正することができる。または、データ主体が直接、裁判所に訴えることもできる。
- 決定内容が変更されず、管理者が法律も遵守している場合、ICOの範疇ではない。ただ、他の機関に訴えることもできる。例えば、ローン審査の場合は、金融オンブズマンなど。

（４）-２　ダイレクトマーケティングのための個人データ処理

- Q4.5：GDPRの第21条2項で「データ主体は、当該ダイレクトマーケティングのための当該データ主体に関する個人データ処理に対して、いつでも異議を唱える権利を持つ」と規定されている。 しかし、そもそもダイレクトマーケティングのための個人データ処理は、（第6条1項(f)の）管理者の正当な利益といった合法性の根拠ではなく、本人の同意（第6条1項(a)）に基づいて行われるべきではないか。換言すれば、ダイレクトマーケティングのケースでは、（第21条2項にいう）データ処理に異議を唱える権利ではなくて、（第7条3項の）同意を撤回する権利を行使するべきではないか？

- A：ダイレクトマーケティングに対する合法性の根拠は、「本人の同意」と「管理者の正当な利益」のどちらも根拠になる。本人同意を根拠とする場合、管理者は同意をいつでも撤回できるということをデータ主体に示す義務を守らないといけない。他方、正当な利益を根拠とする場合、データ主体にはそれを100%拒否する権利がある。
- マーケティングに対しては、他の権利もある。マーケティングの方法によって法律が異なる。これらの法律はEUのePrivacy指令等に基づく。テレマーケティングの場合や電子マーケティングの場合、それぞれ管理者は異なる法律に従う義務がある。ePrivacy指令については改正法案（ePrivacy規則案）の審議中である。

- Q：ダイレクトマーケティングにおいて、どんな場合に「正当な利益」を根拠とできるのか?
- A：データ主体の権利が優先される場合は、「本人同意」が必要。

（４）-３　正当な利益

- Q：ダイレクトマーケティングに限らず、「正当な利益」に関しては、どのような場合に認められるのか、日本でも議論になっている。企業としてはなるべく「本人同意」よりも「正当な利益」を根拠としたいといった意見もある。

- A：「正当な利益」は、企業の利益と個人の権利利益とのバランシングテストが必要であり、決して簡単なオプションではない。企業の「正当な利益」が個人の権利利益よりも優先されるということを企業自体が立証しないといけない。「本人同意」の方がむしろ簡単ではないか。
- ICOでは「合法性の根拠」に関するインタラクティブなツール をWeb公開しており、バランシングテストも行うことができる。
- もし「本人が同意」すれば責任を本人の方に移せるが、「正当な利益」の場合は責任は管理者にある。この場合、ICOに対しても立証責任を負う。

（５）自動意思決定における透明性（１/３）

- Q5.1：自動意思決定のケースにおいて、管理者はどの程度まで「関連するロジックについて意味のある情報」を説明すればよいのか？（特に、自動意思決定がディープラーニングに基づく場合。）

- A：GDPRの透明性のガイドラインや、ICOのガイドもある。「意味のある情報」とは、「アルゴリズムが統計的にどのような変数のパターンを取るとどのような意思決定を行うか」を説明することではない。こうしたことをデータ主体が知ることは、データ主体にとって役立たない。
- 市民が知りたいことは4つある。
  - アルゴリズム（によるプロファイルの作成）にどのようなデータが使われたか？
  - アルゴリズム（によるプロファイルの作成）にそれらのデータがどのように使われたか？
  - どのデータが意思決定に影響を与えたのか？
  - それらのデータがどのような影響を与えたのか？
- 「このデータが変わったら、どのように決定が変わるのか」（反実仮想）を知りたい人もいるので、それを示すことは1つの効果的な説明である。例えば、ローン審査において、良い説明としては以下がある。
  - 収入に対して経費が多すぎる。
  - 追加ローンを申し込むためには、経費を減らせば、6か月後にもう一度見直すことができます。
- すなわち、データ主体の行動をどのように変えれば、よい結果を得ることができるかを説明することである。ICOはこの問題について、アランチューリング・データ科学研究所と協力して枠組みを考えている。

（５）自動意思決定における透明性（２/３）

- Q：GDPRでは事前に一般的な説明を行う義務があるが、データ主体から異議申立があった際など、事後に個別に説明をした方が、良い説明になるのではないか？

- A：次の質問5.2にも関係する。GDPRでは、管理者として前もって正直に理由などを説明することが必要である。「自動意思決定によって、決定のスピードが上がる」など。説明が足りないと思った人は、どこで更なる説明を受けられるか、どこで異議申立（contest）できるかの連絡先を明示する。また、データ源は何か。このアルゴリズムは品質が高いことが保証されている、といったことを事前に説明する。
- 確かに事前の説明は一般的な話になってしまうが、「こういったデータだと、このような決定になる」といった、ある程度の示唆は与えることができる。例えば、オンラインショッピングで注文のボタンを押す前に、「Just-in-time」の通知としてポップアップが表示される。「この情報をローン会社の審査に使います。」同意すると、注文した商品についてクレジットカードの一括払いのみを許すか、分割払いも許すかが自動意思決定される。

- Q5.2：自動意思決定のケースにおいて、処理の透明性を高めるために最も有効な手段は何か？
- A：同上。

（５）自動意思決定における透明性（３/３）

- Q：アルゴリズムの透明性を実現する方法としては、どのようなものがあるか？

- A：アルゴリズムの透明性の方法は色々ある。
- ①反実仮想の説明（Counterfactual Explanations）：

  ローンの申込みに対して単に却下するのではなく、「あなたの年収は2万ポンドですが、年収2万5000ポンドならローンを受けられます。」といった反実仮想の説明を行う。
- ②限定的透明性（Qualified Transparency）：

  市民が理解できなくてもデータ保護監督機関（DPA）が理解できればよい。監査において、AIが予測した12ヶ月分のデータと実際の事実データとを比較して、誰も気づかなかったバイアスが入っていなかったか等を調べる。これにより、透明性、公平性、アカウンタビリティを向上できる。

【ご参考】「反実仮想の説明」とは

- 以下は、Sandra Wachter, Brent Mittelstadt, Chris Russell, 「Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR」（2018年3月更新）の概要である。
- アルゴリズムによる自動意思決定のブラックボックスを開き、一般大衆に分からせようとすることは技術的な障壁が大きい。自動意思決定システムの機能やその原理（rationale）を説明することは技術的にチャレンジングであり、データ主体にとってほとんど意味のない情報であるかもしれない。
- また営業秘密等を含む内容を開示することの法的な障壁もある。アルゴリズムに関する情報は営業秘密を含んだり、その開示によって他者の権利自由（プライバシー等）を侵害したり、データ主体に意思決定の操作を許したりするおそれがある。
- 「自動意思決定に関する説明」は、いかに自動意思決定システムが機能するかについての一般大衆の理解に依存する必要はない。そのような理解可能性が重要であり、追究されるべきであったとしても、「説明」は原則的にブラックボックスを開くことなく提供することができる。データ主体が単に理解することではなく、行為することをサポートする手段として「説明」を捉えるべきである。
- このような「説明」の3つの目的は以下である。
  (1) なぜ個別の意思決定に至ったかを情報提供し、個人が理解することサポートすること
  (2) そのアウトカム（結果）が望ましくない場合は、意思決定に対して異議を唱える基盤（ground）を提供すること
  (3) 現状の意思決定モデルに基づき、将来に望ましい結果を得るために何を変える必要があるかを理解すること
- すなわち、「説明」において、アルゴリズムの「内部の」状態やロジックを伝達するのではなく、当該意思決定を導いた「外部の」事実を記述するべき。「あなたの年収は30000ポンドであるため、ローン審査を通りませんでした。もしあなたの年収が45000ポンドであれば、ローンを受けることができます」といった「反実仮想」（counterfactual）的な説明が有効である。意思決定の結果が望ましくない時に、どうすれば望ましい結果が得られるかの説明と、個人がそれに異議を唱える基盤を与えることが重要である。

（６）機械学習における隠れたバイアス

- Q6.1：機械学習用データにバイアスが含まれている場合、機械学習に基づくプロファイリングは差別を生み出す可能性がある。管理者はこのようなバイアスがかかった機械学習用データをどうやって回避すればよいのか？

- A：これは、GDPRが作られた理由の1つである。データ処理に対するセーフガード（保護措置）が必要となっている。GDPRではシステムの定期的な監査が必要といっている。
  - 1点目は、監査の内容としては、システムが意図された通りに機能しているか。
  - 2点目は、もともとあったバイアスが継続されていないか。
- プライバシーに関するデータ主体の権利のみならず、情報の管理にも関係する話である。
- アルゴリズム自体に対する監査が必要と考える。これはデータ管理者の義務である。これによって、アルゴリズムを継続的に改善することができる。

（７）英国における自動意思決定の実事例

- Q7.1：英国では、GDPR第22条に該当するような本人に重大な影響をもたらす自動意思決定について、実際の事例はあるか？

- A：金融サービスは特に第22条と関わりが深い。詐欺やマネーロンダリングといった犯罪防止のためにデータを使う場合が多い。この場合、色々なデータ源からのデータを使う。
  - 信用格付け会社　　・運転免許庁　　・SNS　など
- 第22条の下に入るケースか否かの判断は難しい。
  - 「人が関与していない（完全に自動化）」かどうか。
  - 「本人に法的効果または同様の重大な影響をもたらす」かどうか。
- 最近、自動意思決定は様々な分野で広がりを見せている。例えば、警察が容疑者を検挙した際、容疑者をすぐに裁判にかける方がよいのか、それとも地域で保護観察期間を置く方がよいのかを決めるのに、ツールを使っている。現在は警察官が最終的に判断しているが、将来的には大部分を機械で行うことがあることは想像しやすい。このツール（システム）を使うにはリスクアセスメントが必要である。すなわち、すぐに刑務所に入れることのリスクと、保護観察期間を置くことによる再犯リスクがあり、どこにリスクを置くのかという判断が必要になる。

- Q：EUではなぜ、プロファイリングや自動意思決定に対する懸念が大きいのか。何か歴史的な背景や、きっかけとなる大きな事件があったのか？
- A：テレマティクス保険（リアルタイムに提供する運転情報をもとに保険料を算出する。英国や米国では導入が進んでいる）の問題などが影響している。消費者団体などがデータ保護の立場から、きちんとした判断が必要と言っている。

（８）プロファイリングと自律性

- Q8.1：本人に重大な影響をもたらさないとしても、プロファイリングは本人の自律性を侵害する（自律的な意思決定を妨げる）可能性がある。例えば、ターゲティング広告やパーソナライズされたウェブコンテンツは、いわゆる「フィルターバブル」を生み出す恐れがある。管理者や本人は、このように自律性が侵害されることをどのように回避すればよいか。

- A：EUでも「フィルターバブル」の話は沢山している。個人の世界観が狭くなる。透明性の問題もある。まず、個人個人がなぜ、自分の所にこんなコンテンツが来るのかを理解するべき。そして、設定を変える方法を知るべき。一部のソーシャルメディアではこのような努力をしている。Facebookやtwitterでは、透明性を挙げるための努力がなされている。

（９）プロファイリングの正確性（１/２）

- Q：AIの判断の正確性がまだ足りないのでは？

- A：AIはまだ幼稚な面がある。データがクリーンかどうかを確かめなければならない。バイアスが含まれているかどうかを調べる時点で、その人自身が気付かないと、入力データにバイアスが入ってしまう。もともと持っているバイアス、無意識的なバイアスを知らないと、クリーンなデータにならないので、大きな課題である。
- もう1つの問題として、AIのプロファイリングでは、データとデータの相関関係に基づいて推測（inference）したり、自動意思決定したりする。相関関係があいまいなので、データと決定との関係（入力データと出力データの関係）が明らかでない場合が多い。例えば、AIで自動車保険の料金を決める場合、ソーシャルメディアのデータを用い、書き込み内容ではなく、書き方や表現方法だけに注目したケースがある。すなわち、「！」の数やエキサイトした書き方のみで、自動車事故のリスクが高いと決定していた。このアルゴリズムは実際にはパイロットのみで、実際には使われなかった。今の段階では、相関関係があると出てしまうかもしれない。長期的には、相関関係がなかったと分かるかもしれない。

- Q：ブラックボックス型のAIのみならず、ホワイトボックス型のAIが増えればよいのではないか？

- A：人の生命にかかわる危険な分野もある。ブラックボックスAIが何を目的として作られたものかによる。例えば、医療で補助的な使われ方をするのであれば、非常に役に立つ。逆に、AIでの判断結果によって、自分の金銭的側面に影響があると嫌だなと思う。ブラックボックスの改善はゆっくりと進めるべき。
- 囲碁のAI（アルファGo）では、それぞれの手をどうしてそのように打ったのかの根拠は分からない。囲碁ならばよいが、他の分野ではAIの判断の理由や根拠を理解することは重要である。

（９）プロファイリングの正確性（２/２）

- Q：先ほどの自動車保険の事例で、仮に「！」の数と自動車事故のリスクの高さの相関関係が高いということが長期的に分かってしまったとしたら、我々はそのような根拠も受け入れなければならないのか？

- A：システムが導き出す決定や相関関係は、慎重に見なければならない。例えば、Google Fluの事例では、1年目、2年目は相関があるとみられたが、3年目は相関がないことが分かった。
- EUでは、データの正確性について議論がある。例えば、病院でAIが、「カールさんは心臓発作を起こす確率が高い」と判断した。しかし、結局カールさんが心臓発作を起こすことはなかった。この場合、AIの判断は間違っていたのか？GDPRではデータは正確でないといけないとされ、不正確な場合は本人が訂正する権利がある。心臓発作を起こさなかったにも関わらず、AIの判断が正しかったとすれば、本人に訂正する権利がない。何が正しいか。このケースでは、AIが、カールさんを集団（Population）の中で「心臓発作の可能性が高い」というカテゴリーに分類したのは正しいだろう。

- Q：「心臓発作の確率が高い」とプロファイリングされたことで保険料が上がっていたとしたら、（心臓発作がその後起こらなかったならば）不正確なデータとして本人は訂正の権利を行使したい。しかし、今後AIが進歩してAIの判断が100%近い精度のものになると、訂正は難しくなるのではないか。

- A：AIの判断の要素（factor）を一層理解しにくくなるかもしれない。その場合、人間の介在が必要だろう。この分野は、EU、英国のみならず、国際的な協調が必要である。