

2017年12月20日

情報通信法学研究会（新領域分科会）

プロファイリングの法的諸論点(試論)——憲法の観点から

山本龍彦（慶應義塾大学）

I. はじめに

(1)「プロファイリング」の定義

・EU「一般データ保護規則（General Data Protection Regulation, 以下「GDPA」という。）」4条4項：「自然人に関する特定の個人的側面を評価するために、特に、当該自然人の職務遂行能力、経済状況、健康、個人的選好、関心、信頼性、行動、位置もしくは動向を分析（analyse）または予測（predict）するために、個人データを用いて行うあらゆる形式の自動化された個人データ処理」

・プロファイリング3要素（「個人に関する自動化された意思決定およびプロファイリングに関するガイドライン（*Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/678*, Adopted on October 2017. 以下「ガイドライン」という。）」）

①自動化された処理形式

②パーソナル・データに関する処理

③自然人の 個人的側面（personal aspects） を評価するという目的（以上、6頁）

※1 特に、個人的側面の「分析または予測」を目的として行われるものを指すが、「年齢や性別のような特性に基づいて個人を単に査定（assessing）または分類（classifying）することも、予測的目的にかかわらず、プロファイリングとみなされうる」（7頁）

(2)プロファイリングの実例

①米国ターゲット社の事例（妊娠予測）

②新卒採用

※2 ソフトバンクが新卒採用に IBM 社の AI である Watson を使用することについて、https://headlines.yahoo.co.jp/hl?a=20170529-00000035-zdn_ep-sci

③個人向け融資の判断

※3 「個人融資、AI 審査」日本経済新聞朝刊平成 29 年 5 月 28 日 5 面（「みずほ銀行とソフトバンクは 9 月から、人工知能（AI）を使った個人向け融資を始める。銀行口座の入出金履歴や携帯電話の利用料金の支払い状況などをもとに信用力を点数化し、最短 30 分で融資する」）。

※4 インステックにおける利用（AIG は AI プロファイリングを用いたリスク査定の効率化を目指す。日本経済新聞朝刊平成 29 年 9 月 6 日）。

④米国各州における再犯リスク評価

※5 *State v. Loomis*, 881 N. W. 2d 749(Wis. 2016) : “COMPAS” というシステムにより再犯リスクをスコアリングされ、裁判所に同スコアを踏まえて量刑判断された被告人（アフリカ系アメリカ人）が、正確性が担保されず、検証可能性もない同システムによりスコアリングされることは不公正であり、デュー・プロセスに反するなど主張した事件。

※6 COMPAS : Northpointe 社の構築した再犯リスク評価システム。正式名称は、“Correctional Offender Management Profiling for Alternative Sanctions（代替的制裁のための矯正の犯罪者管理プロファイリング）”。COMPASは、犯罪歴、雇用状況、教育レベル、家族の犯罪歴、信条を含む130以上の情報から、再犯リスクを10段階で評価する。

※7 裁判官の量刑判断の際にプロファイリングが用いられている州として、アリゾナ、コロラド、デラウェア、ケンタッキー、ルイジアナ、オクラホマ、ヴァージニア、ワシントン、ウィスコンシンがある。Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, *ProPublica, Machine Bias*, ProPublica, May 23, 2016, at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

⑤Facebook を用いた実証実験

※8 Facebook の画面上に投票所の場所、既に投票を済ませた友人の写真、自らが投票したことを友人に告げるボタンを選択的に表示された者は、そうでない者と比して実際に投票に行く可能性が高くなることを実証。そうすると、政治的プロファイリングによりある特定政党を支持しようと予測された者にのみこうした情報を表示・フィードすれば、選挙結果の操作も可能となる（「デジタル・ゲリマンダリング（digital gerrymandering）」）。Jonathan Zittrain, *Engineering an Election*, 127 HARV. L. REV. F. 335, 336-336(2014).

(3)「憲法」問題の概観

・プライバシーの「迂回的」侵害

・社会的排除・差別（バーチャル・スラム）——個人の尊重

・国民主権・民主主義の歪み？

→「プロファイリングと自動化された意思決定は、個人の諸権利と諸自由と重大なリスク（significant risks）を与え得る。それは、適切なセーフガードを要求するものである」（ガイドライン5頁）。

Ⅱ.「現代的プロファイリング」は、古典的なプロファイリングと何が違うのか？

①データ量（スモールデータ／ビッグデータ）

②自動性（人間の経験則・勘／アルゴリズムによる自動的評価）

③科学的信憑性・外見的客観性(based on 豊富なデータ量と人間の直感を排した自動処理)

→余剰存在による訂正容易性（古）／余剰縮減による訂正困難性（現）

④意外性（予見困難性）

1 ※2 1人当たりのチーズの消費量と、ベットシーツに絡まって死亡する者の数との相関関係。ビールとお
2 むつ。カーリーポテトと知性。人間が予見し難い意外性の高いデータがプロファイリングに広く用いられ
3 る傾向。

4 →萎縮効果（何が評価対象にあるかわからない）

5 ⑤項目の広汎性・細目性

6 →AI がますます「賢く」なっていけば、これまで人間が行ってきた重要な決定の多くを AI
7 （現代的プロファイリング）が代替するようになるだろう。

8 ※3 程度の相違論（本質的相違否定論）への批判（住基ネット判決、GPS 判決）

9

10 **Ⅲ.「現代的プロファイリング」の「憲法」問題**

11

12 **1. プライバシー権の侵害リスク(憲法 13 条)——迂回的「取得」問題**

13

14 **(1)問題の所在**

15 →取得に厳格な同意が要求されるようなセンシティブ情報（要配慮個人情報。個情法 17 条
16 2 項）も、周知的な情報を分析（プロファイリング）させることで、迂回的に「取得」（厳
17 密には「推知」）できてしまう（妊娠、鬱病、政治的信条 etc.）。

18 →「消費者としては心の奥底に秘めた私事も、AI には常に見透かされていることを覚悟す
19 る必要がある」（山本）

20 ※1 個情法 17 条 2 項：「個人情報取扱事業者は、……あらかじめ本人同意を得ないで、要配慮個人情報を
21 取得してはならない」

22

23 **(2)取得・推知論争**

24 ※2 私見は、AI を用いた一定のプロファイリングについて、取得同視論。

25 ※3 「一般の人が……当該私人の私生活であると誤認しても不合理でない程度に真実らしく受け取られる
26 ものであれば、それはなおプライバシーの侵害としてとらえることができる」（宴のあと事件判決。東京地
27 判昭和 39 年 9 月 28 日下民集 15 卷 9 号 2317 頁）

28

29 →私見につき、宇賀克也『個人情報保護法の逐条解説〔第 5 版〕』（有斐閣、2016 年）132
30 頁は、「本人の同意なしにプロファイリングによって要配慮個人情報を新たに生み出すこと
31 は、要配慮個人情報を新たに『取得』に当たると解すべきかどうかという重要な解釈問題
32 が存在する」と述べる。また、一般に取得否定論をとる藤原静雄も、私見につき、「憲法学
33 の立場からの貴重な指摘であり、傾聴に値する」とし、「一般的には難しい」ものの、「こと
34 信条となるとそのような解釈の余地は否定しない方がいい場合がある」としている（宇賀克
35 也ほか「〔鼎談〕個人情報保護法改正の意義と課題」行政法研究 13 号（2016 年）11 頁〔藤
36 原発言〕）。

1 ※4 平成 29 年 5 月 30 日衆議院憲法調査会・宍戸常寿参考人発言も参照。なお、宍戸が策定に関わった
2 「放送受信者等の個人情報保護に関するガイドライン」（平成 29 年 4 月 27 日総務省告示第 159 号）の 34
3 条は、「受信者情報取扱事業者は、視聴履歴を取り扱うに当たっては、要配慮個人情報を推知し、又は第三
4 者に推知させることのないよう注意しなければならない」と規定し、プロファイリングにより視聴履歴か
5 ら要配慮個人情報を推知することを制限した。この規定は日本で最初のプロファイリング規制といえる。
6 ※5 「プロファイリングのプロセスは、データ主体にとってしばしば不可視である。それは、個人に関する
7 派生的または推論的データ——データ主体自身から直接提供されていない『新たな』データである——
8 を創造することによってなされる」（ガイドライン 17 頁）。

10 2. 平等原則(憲法 14 条)、個人の尊重原理(憲法 13 条)の侵害リスク

12 (1)歴史的な被差別集団に対する差別・排除の再生産

13 →2016 年の FTC レポート（『Big Data: A Tool for Inclusion or Exclusion?』）など

14 ※1 他にも、大統領府が 2016 年 5 月に公表した『ビッグデータ：アルゴリズムックシステム、機会、市
15 民権に関する報告書』が重要である。EXECUTIVE OFFICE OF THE PRESIDENT, BIG DATA: A REPORT
16 ON ALGORITHMIC SYSTEMS, OPPORTUNITY, AND CIVIL RIGHTS(MAY 2016).

17 →米国では特に意識される（人種的マイノリティへの差別的インパクト etc.）

19 (i) なぜそれが起きるか？

20 ①過少代表（underrepresentation）

21 ※2 アメリカのボストン市は、道路状況の調査のため市民のスマートフォンから得られる GPS データを
22 利用したが、その途中で、これだと高所得者の居住エリアに道路補修サービスが集中してしまうことが判
23 明した。低所得者のなかにはスマートフォンを所持していない者が多く、低所得者の居住エリアからのデ
24 ータが解析母数に過少に代表されてしまったためである（データの偏りとマイノリティ）

25 ※3 プライバシー信奉者のマイノリティ化（以下（2）へ）

26 ②既存バイアスの反映

27 ※4 例えば、現在の従業員データを使って、「良い従業員」を予測するアルゴリズムを構築した際、現在
28 職場に存在しているバイアスをアルゴリズムが承継してしまうことになる。実際、イギリスの聖ジョージ
29 病院は、人種的マイノリティと女性に対して不利のあった従前の入学試験データに基づいて「良い医学生」
30 を発見するアルゴリズムを構築したために、この選考によって同様のバイアスを再生産することになった
31 という

33 (ii) アメリカの対応

34 →FTC レポートは、アルゴリズム予測の利用者は、以下に掲げる 4 つの事項を考慮すべき
35 であると述べている。

36 ①利用するデータセットがあらゆるコミュニティを適切に代表したものになっているか

- 1 (特定の人々からの情報を特に欠いているということはないか)、
2 ②隠れたるバイアス (hidden biases) が特定の人々に意図せざるインパクトを及ぼしてい
3 るということはないか、
4 ③「うわべだけの相関関係」がアルゴリズムに組み込まれていないか (医療、与信、雇用な
5 どの重要な決定 [important decisions] にアルゴリズムが用いられる場合には、アルゴリズ
6 ムの適切さに対する人間のチェックが確保されているか)、
7 ④その利用に当たって公正さが考慮されているか (予測精度と公正との衡量)。

8 ※5 日本でも、総務省設置のAIネットワーク社会推進会議が2017年7月に公表した『国際的な議論の
9 ためのAI開発ガイドライン案』において、AIを用いたプロファイリングによる差別などの問題を念頭
10 に置いて、「開発者は、採用する技術の特性に照らし可能な範囲で、AIシステムの学習データに含まれる
11 偏見などに起因して不当な差別が生じないよう所要の措置を講ずるよう努めることが望ましい」と述べら
12 れている。

13 14 (2)新たな被差別集団の形成——バーチャル・スラム

15 【シナリオA】通信事業を行うY社は、新卒採用に、採用希望者の適性をAIに予測評価させる
16 プロファイリングシステムを導入している。ただし、この予測評価を導くアルゴリズムは公開し
17 ておらず、希望者のどのような情報を、どのようなウェイトで考慮しているのか、外部には具体
18 的にわからないようにしている。

19 大学4年生のXは、Y社への入社を希望し、エントリーシートを送付したが、その後、不採用
20 を知らせる連絡を受けた。Xは、エントリーシートに記入した事項以外の情報 (例えば、SNSに
21 関する情報) も広くAIの予測評価に使われていることを知り、自分の何がAIの評価を下げた
22 のだろうと悩み始めた (SNSで市民運動家のAと「友達」になっているの悪かったのか、数年
23 前に外国の反政府活動家の写真に「いいね」を押したのが悪かったのか)。どうしてもY社に入
24 社したいXは、就職浪人することに決め、卒業後にファスト・フード店でアルバイトを始めた。
25 しかし、翌年も、エントリーシートの送付後に不採用を知らせる連絡を受けた。Xは、ために、
26 Y社のプロファイリングシステムと同じシステムを導入しているB社やC社にもエントリーシ
27 ートを送付したが、やはり不採用の連絡を受けた。

28 Xは、思い切って、Y社に対して、AIの「意思決定」プロセスを開示するよう求めたが、営
29 業上の秘密に当たるなどとして拒否された。Xは、その後もAIに「嫌われる」理由がわからず、
30 自己改善の方向性もわからぬまま、採用にAIによるプロファイリングシステムを導入していな
31 い低賃金のアルバイト職を転々とした。

32 Xは、アルバイトをしながら独学でビジネスの勉強を始め、30歳を迎えたときに自ら事業を
33 始めようと、D銀行に対してオンライン上で融資を申し込んだ。D銀行は、融資判断の際に、申
34 込者の返済率や信用力をAIに予測評価させるプロファイリングシステムを導入している。Xは
35 嫌な予感がしたが、案の定、オンライン上で、D銀行から融資できないとの回答を受けた。そ
36 ども、予測評価のアルゴリズムや、融資が拒否された理由などが開示されることはなかった。そ

1 のため X は、融資が拒否された理由が、低賃金のアルバイトを長年続けていたことにあるのか、
2 それとも別のところにあるのかがわからず、途方に暮れることとなった。

3 X は、その後数年、AI の予測評価システムを導入しているあらゆる組織から排除され続け、
4 自らが社会的に劣った存在であると感じるようになった。X は、自分以外にも、明確な理由もわ
5 ならず AI の予測評価によって社会的に排除され続けている者たちが多数存在し、仮想空間にお
6 いて「スラム」(バーチャル・スラム) を形成していることを知った。しかし X は、その者たち
7 と SNS で交流すると、信用力などに関する AI の予測評価がさらに下がるという噂を聞いてい
8 たために、悩みを共有する相手もないまま、孤立を深めていた。そのとき、携帯していたスマ
9 ートフォンが鳴り、日々の生活記録などからユーザーの健康を管理する団体から、「あなたは
10 ま鬱状態にあるようです。気を付けてくださいね」というメッセージが届いた。X は、そのメッ
11 セージをみて、かえって生きる気力を失った。

12 ※1 Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated*
13 *Predictions*, 89 WASH. L. REV. 1, 32(2014)をベースに作成。

14
15 →「確率という名の牢獄」(マイヤー＝ショーンベルガー)、「負のスパイラル」(Danielle
16 Keats Citron & Frank Pasquale)、「バーチャル・スラム」(山本)

17 18 (i)セグメント主義と個人の尊重

19 →プロファイリング結果は、あくまでもビッグデータに基づく確率的な判断であって、その人
20 そのものの確定記述ではない(個別化/個人化)。が、その科学的信憑性から、プロファイ
21 ング結果があたかもその人の「実像」であるかのように扱われる可能性(自動化バイアス)、
22 あるいは、実像そのものでない理解されても、経済的合理性や効率性の観点から、人生の
23 重要場面で当該評価がその人に関する最終的評価として扱われる可能性がある(セグメン
24 トに基づく類型的、概括的、事前的判断)。

25 →ベルトコンベアに載せられ、仕分けされる個人

26 ※2 「個人の尊重」の近代的意義との緊張

27 →前近代＝身分のような集团的属性により個人が短絡的・概括的に判断されたり、生き方が事前に規定さ
28 れたりする世界 (but 効率性の世界)

29 →近代＝一人ひとりの具体的事情を尊重して(時間とコストをかけて一人ひとりをみて)、個人が主体的・
30 自律的に人生を創造することを認める世界(個人の尊重)

31 →「個人の人格や能力を AI により確率的に判断し、様々な可能性を事前に否定することは個人の尊重原
32 理と鋭く矛盾する。最先端の情報技術を使って、われわれを固定的で予定調和的な『前近代的』世界へと
33 引き戻すようなものだ」(山本)。

34 35 (ii)ブラックボックス化と個人の自律

36 →また、アルゴリズムの高度の専門性・複雑性や不透明性から、個人がプロファイリング結

- 1 果の妥当性を自ら検証し、これに異議を差し挟むことも現実に困難となる。理由もわからぬ
2 まま（何を、どのように評価付けされたのかもわからぬまま）、また、それゆえに自己改善
3 の方向性もわからぬまま、ただただ没落していく。
- 4 →ブラックボックス化による自己実現の困難（自律的に生きることの困難性）
- 5 ※3 不条理な没落（カフカ的世界）：反論の機会？
- 6 ※4 演技（ペルソナ）の失効（近代的人格概念）
- 7 ※5 「自然」と「作為」という峻別の失効（丸山眞男）
- 8 →後にガイドラインを踏まえて検討
- 9
- 10 (iii)“More Data”原理と個人の尊重——のしかかる過去
- 11 ・自然的傾向としての“More Data”原理（AIの予測精度とデータ量との比例関係）
- 12
- 13 ・ノンフィクション『逆転』判決（最判平6・2・8）
- 14 →「更生を妨げられない利益」（人生をやり直す自由？）
- 15 →データ・スティグマから解放する必要性（忘れられる権利？、「越境」の自由）
- 16 ※5 「ある者が刑事事件につき被疑者とされ、さらには被告人として公訴を提起されて判決を受け、とり
17 わけ有罪判決を受け、服役したという事実は、その者の名誉あるいは信用に直接にかかわる事項であるから、その者は、みだりに右の前科等にかかわる事実を公表されないことにつき、法的保護に値する利益を
18 有するものというべきである。「この理は、右の前科等にかかわる事実の公表が公的機関によるものであ
19 っても、私人又は私的団体によるものであっても変わるものではない。そして、その者が有罪判決を受け
20 た後あるいは服役を終えた後においては、一市民として社会に復帰することが期待されるのであるから、
21 その者は、前科等にかかわる事実の公表によって、新しく形成している社会生活の平穏を害されその更生
22 を妨げられない利益を有するというべきである」。
- 23
- 24 「本件事件及び本件裁判から本件著作が刊行されるまでに12年余の歳月を経過しているが、その間、被上
25 告人が社会復帰に努め、新たな生活環境を形成していた事実に照らせば、被上告人は、その前科にかかわ
26 る事実を公表されないことにつき法的保護に値する利益を有していたことは明らかである」。「しかも、被
27 上告人は、地元を離れて大都会の中で無名の一市民として生活していたのであって、公的立場にある人物
28 のようにその社会的活動に対する批判ないし評価の一資料として前科にかかわる事実の公表を受忍しなけ
29 ればならない場合ではない」。
- 30
- 31 ・相続分規定違憲決定（最大決平25・9・4）
- 32 →「子を個人として尊重〔すべき〕」という考え方から、「〔子が〕自ら選択ないし修正する
33 余地のない事柄を理由としてその子に不利益を及ぼすことは許され〔ない〕」の述べ、同規
34 定を違憲
- 35
- 36 →プロファイリングに原則として用いてはならない情報を括り出すべきか？ それは何か？

1 (アルゴリズム上の操作の必要?)

2 ※6 GDPR22条4項(特別類型に属する個人データを含む自動的意思決定を原則として禁止。ただし、9
3 条2項(a)および(g)による例外)

5 3. 民主主義原理(国民主権原理)の侵害リスク

7 ・プロファイリングによる「フィルター・バブル」と集団分極化
8 →社会的分断と民主主義

9 ※1 選挙権の適切な行使ないし選挙の公正の確保には、有権者に政党や候補者に関する情報が「適正に」
10 あるいは「自由かつ均等に」伝達されることが必要(最大判平成19・6・13民集61巻4号1617頁〔田原
11 反対意見、泉反対意見〕)。

13 ・政治的信条のプロファイリングと選挙

14 →デジタル・ゲリマンダリングと「選挙の公正」、選挙の「マーケティング化」(熟議過程の
15 喪失と統治の正統性危機)

16 ※2 政治的プロファイリングを用いた操作的な選挙過程は否定されるべき

17 ※3 戸別訪問の禁止の意義(公選法)

20 IV. GDPRによる現代的プロファイリングの規制と諸権利——プロファイリングによる 21 確率的判断への対抗?

23 1. 異議を唱える権利(中止請求権)

25 →GDPRは、データ主体(以下、「主体」という。)に対し、プロファイリングに対して異
26 議を唱える権利(right to object)を認めた。同権利が行使された場合、データ管理者(以
27 下、「管理者」という。)は、主体の利益等を乗り越える「やむにやまれぬ正統な根拠
28 (compelling legitimate grounds)」を示さない限り、プロファイリングを中止しなければ
29 ならない(21条1項)。また、管理者は、主体による権利行使を可能にするために、主体が
30 当該権利を持つことを明確なかたちで告知しなければならない(21条4項)。

31 ※1 なお、ダイレクト・マーケティングと関連したプロファイリングに対し同権利が行使された場合、管
32 理者はいかなる場合でもプロファイリングを中止しなければならない(21条2、3項)。

33 ※2 オプトアウト的な方向性?

34 ※3 「GDPRは、どのようなものがやむにやまれぬ正統な根拠とみなされるかについていかなる説明を提
35 供していない」。「管理者の単なるビジネス上の利益」がこれに当たらない一方で、「伝染病の蔓延を予測す
36 るためのプロファイリング」はこれに当たり得る。「常に管理者のやまぬやまれぬ利益と、データ主体の異

1 議の根拠との比較衡量 (balancing) が求められる」。ただ、「やまぬやまれぬ正統な根拠を証明する責任は、
2 データ主体ではなく、管理者に置かれる」(ガイドライン 25 頁)。

3 ※4 なお、GDPR16 条の訂正権、17 条の消去権は、「インプットデータ」(プロフィールを作成するため
4 に使われる個人データ) とともに、「アウトプットデータ」(「プロフィールそれ自体または個人に割り当て
5 られた『スコア』」にも適用される (ガイドライン 24 - 25 頁)。

7 2. 自動処理のみに基づき重要な決定を下されない権利

9 GDPR22 条:「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
10 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取扱いのみに
11 基づいた決定に服しない権利を持つ」(The data subject shall have the right not to be
12 subject to a decision based solely on automated processing, including profiling, which
13 produces legal effects concerning him or her or similarly significantly affects him or her)

15 ※1 EUは、1995 年プライバシー指令 (Data Protection Directive) の時代から、「個人の特徴に関する、
16 完全に自動化された評価は、当該個人の利益を大きく侵害するような決定の唯一の根拠となつてはならな
17 い」という考えを個人情報保護法制の核心としたきた。GDPR では、ビッグデータの利活用やAI 技術の
18 進展を見込んで、こうした核心的思想を明確にしたうえで、規制を拡充したものと解することができる (Lee
19 A. Bygrave, *Minding the Machine: Article 15 of the EC Data Protection Directive and Automated*
20 *Profiling*, 17 COMPUTER LAW & SECURITY REVIEW 17, 24(2001).)。

22 (1)プロファイリングの 3 層(ガイドライン 8 頁)

23 ①一般的プロファイリング

24 ②プロファイリングに基づく意思決定 (例えば、「純粋に自動的な手段によって生み出され
25 たプロフィールに基づいてローンに同意するかどうかを 人間が決める」)

26 ③自動処理のみに基づく意思決定 (「有意味な人間のインプットなしで、ローンに同意する
27 かどうか、決定が自動的に個人に送られるかどうかを アルゴリズムが決める」)

28 →22 条の射程は③

30 →ここで、「のみに基づく」とは、「決定プロセスにいかなる人間の関与も存在しない」こと
31 を意味するが、「管理者は、人間の関与をでっち上げる (fabricating) ことにより 22 条を潜
32 脱することはできない」。例えば、「結果への実際上の影響 [を考慮すること] なしで、自動
33 的に生成されたプロフィールを決まって (routinely) 個人に適用するような場合、それは、
34 自動化された処理のみに基づく決定としてみなされる」。また、「人間の関与として認められ
35 るに、管理者は、決定の監督が単なる名ばかりのジェスチャー (token gesture) ではなく、
36 実質的なものであることを保証しなければならない。それは、決定を変更する権限と能力を

1 持つ者によって実行されるべきである。分析の一部として、彼らは、すべての利用可能なイ
2 ンプットデータおよびアウトプットデータを考慮すべきである」(ガイドライン9 - 10頁)。
3 →人間の^{実質的}関与を要求(「人間の関与は鍵となる要素 (key element) である」〔ガイド
4 ライン15頁〕)
5 =人間関与原則

6

7 (2)「法的な効果」の意味

8 →「他者と関係 (associate) する自由、選挙で投票する自由、訴訟を提起する自由のような、
9 誰かの法的権利に影響する处理的活動」。「法的な影響は、個人の法的地位 (legal status)
10 や契約の下での権利に影響する何かともいえる」。

11 例えば、

- 12 ①法によって付与される特定の社会保障の資格が認められたり、否定されたりすること
- 13 ②国境への立入りを拒絶されること
- 14 ③当局によって、高められたセキュリティ手段または監視に服すること (以上、ガイドライ
15 ン10頁)

16

17 (3)「同様に重大な影響」の意味

18 →「個人の状況、振る舞い、選択に重大な影響を与える可能性をもつ」もの

- 19 ・排除や差別をもたらす決定
- 20 ・オンラインのクレジット申請の自動拒否 (リサイタル71)
- 21 ・いかなる人間的関与もない電子的な採用活動 (同)
- 22 ・一定のターゲティング広告

23 ※1 「多くの典型的なケースにおいて、ターゲティング広告は、個人に重大な効果を与えるものではない。
24 「しかしながら、ケースの具体的性質によって、そのように考えられることはあり得る」(ガイドライン11
25 頁)。

26 ①プロファイリング過程の侵害性

27 ②個人の期待や願望

28 ③広告が送られる方法

29 ④標的とされたデータ主体の脆弱性 (particular vulnerabilities)

30 「一般には、個人に対してわずかな影響しか与えない処理が、社会における特定集団に重大な効果を実際
31 上与えることがあり得る。例えば、財政的な困難に陥っている者で、オンラインのギャンブルの広告を規
32 則的に見せられた者は、こうした申し出にサインし、さらなる負債を負う可能性がある」

33

34 【シナリオB】メタボリック症候群を気にし始めた30代男性であるBは、ネット上でダイエット関連の商
35 品をチェックしたり、実際にトレーニング・ジムを見学したりしていた。

36 ダイエット食品を扱う健康食品会社Cは、Bのウェブ閲覧記録や、行動履歴(GPS位置情報)などを

1 継続的・網羅的に収集しており、ビッグデータ解析と、これらの情報を用いたプロファイリングの結果か
2 ら、Bがメタボを気にする30代男性であると予測していた。

3 そこでC社は、ニュース配信などのサービスを行うインターネット広告事業者Dに、同社のポータルサ
4 イト上で、一般的なニュースとともに、メタボの健康リスクを報じるニュースを、Bに対してのみ数日間
5 にわたって集中的に配信するよう依頼した。

6 その後、C社は、Bがこのニュースを読み、メタボへの不安をさらに増大させていることを他の行動履
7 歴などから確認できたため、そのタイミングでBに対してダイエット食品のネット広告を配信した。

8 山本龍彦『おそろしいビッグデータ』（朝日新聞出版、2017年）

9
10 ※2 「例えば、ひどく高い値段が、事実上、誰かが特定の商品やサービスを得ることを禁止するような場
11 合には、異なる値段をつけるような自動的意思決定は重大な効果を持ち得る」（ガイドライン11頁）。

12 ※3 重大な効果は、ポジティブな場合もネガティブな場合もあり得る（同）。

13 14 (3)例外

15 →GDPRは、

16 ①主体と管理者間の契約の実現または締結に必要な場合（ただし、「この必要性は、限定的
17 に解釈されるべき」とされている。ガイドライン12頁）、

18 ②EUまたは加盟国の法によって承認されている場合（例えば、管理者が提供するサービス
19 の安全や信頼性を保証するために、詐欺や脱税を監督し、予防するような自動的意思決定を
20 認める立法があり得る。ガイドライン12頁）、

21 ③主体の明示的な同意に基づく場合には、プロファイリングのみに基づく重要決定もなさ
22 れうるとしている（22条2項）。

23 もっとも、この場合でも、管理者は、主体の権利・自由等を保護するために適切な措置を
24 講じなければならず、少なくとも人間の介入を得る権利（right to obtain human
25 intervention）、自らの見解を表明する権利、決定を争う権利を保障しなければならないと
26 される（22条3項）。

27 28 3. 透明性の要請

29 30 (1)告知の内容

31 →GDPRは、「公正と透明性を確保するために」、管理者は、①こうした決定の存在はもち
32 ろん、②当該決定の「ロジックに関する意味のある情報」、③「その処理の重大性およびデ
33 ータ主体に及ぼす想定される帰結」を主体に対して告知しなければならないとしている（13
34 条2項(f)）。他方で、主体は、このような情報にアクセスする権利を有するものとされる（15
35 条1項(h)）。また、GDPRは、上述のような決定がなされる場合に、データ保護影響評
36 価の実施を「特に（in particular）」求めている（35条3項）。

1
2 →22 条の場合に、透明性義務 (transparency obligations) がとくに高められた (ガイドラ
3 イン 13 頁)。

4 ※1 「ブラックボックス化」に対する懸念の解消

6 (2)「ロジックに関する意味のある情報(meaningful information about logic involved)」

7
8 →「機械学習の発展と複雑性は、自動化された意思決定プロセスやプロファイリングがどの
9 ように行われているのかを理解することを難しくしている」。「管理者は、用いられるアルゴ
10 リズムの複雑な説明を試みたり、完全なアルゴリズムの開示したりすることなく、当該決定
11 にたどりつくうえで依拠した基準や、その背景的な根拠についてデータ主体に告知する シ
12 ンプルな方法」を探すべきである。提供される情報は、データ主体にとって有意味なもので
13 あるべきである」(ガイドライン 14 頁)。

14 →完全なアルゴリズムの開示までは求められない。

15
16 例えば、

17 ・「決定にたどりつくうえで考慮された主な特性の詳細」

18 ・「情報の源泉と重要性」(どこから情報を得たのか)

19 ※2 もっとも、GDPR リサイタル 63 は、商業上のセンシティブ情報に対して一定の保護を与えている(情
20 報へのアクセス権は、取引の秘密または知的財産、とくにソフトウェアを保護する著作権を含む、他者の
21 権利または自由に不利な影響を与えてはならない)。

23 (3)「処理の重大性と想定される帰結」

24
25 「保険会社は、顧客の運転時の振る舞いに基づき、自動車保険の保険料を設定するために自
26 動化された意思決定プロセスを用いる。この処理の重大性および想定される帰結を描くた
27 めに、会社は、危険な運転が、高い保険料支払いの原因となることを説明し、アクセルのふ
28 かし癖や直前のブレーキングのような危険な傾向をもった架空の運転者と比較するアプリ
29 を提供する」。「保険会社は、どうすればこうした癖を改善でき、どうすれば保険料を下げら
30 れるのかを示すヒントとなるようなグラフィックスを用いる」(ガイドライン 15 頁)。

31 ※3 個人の自律への配慮(ホワイトボックス化) = 人生をやり直す自由

33 5. アメリカにおける対応

34
35 ・ State v. Loomis, 881 N. W. 2d 749(Wis. 2016)

36 →ウィスコンシン州最高裁判所は、量刑判断において COMPAS の使用を認める判断を下し

1 た。しかし、重要なのは、州最高裁が、裁判所による同システムの利用に以下のような条件
2 を付したことである。

3 ①同システムによるリスク・スコアは、裁判官により最終的なものとして考慮されてはなら
4 ない（依拠／考慮）、

5 ②デュー・プロセスを確保するという観点から、裁判官に対し以下のことが警告されなけれ
6 ばならない。

7 (a) COMPAS の知的財産的性格から、諸要因がいかんにかに衡量され、リスク・スコアがどの
8 ように導出されるのかに関する情報は開示されないこと、

9 (b) リスク評価の有効性について、ウィスコンシン州の住民のみを前提とした調査はい
10 まだ完了していないこと、

11 (c) いくつかの研究調査では、再犯リスクが高いとして、人種的少数派に属する犯罪者
12 が不当に分類されることについて懸念が表明されていること、

13 ※1 判例法としてのデュー・プロセス。ただし、*Recent Cases*, 130 HARV. L. REV. 1530(2017)の痛烈
14 な批判。裁判官はアルゴリズムによるリスク評価の妥当性を査定する能力をもたないうえ、かかるリスク
15 評価を用いるべきとの強い圧力を受けるために、結局このリスク評価に「依存」することになると指摘。
16 すなわち、裁判官は、かかる評価を正當に「考慮」して自律的に量刑判断するために必要なアルゴリズム
17 への「懐疑」を持つことができない（違憲状態の示唆）。

18 19 ・ FTC レポート

20 →公正信用報告法（Fair Credit Reporting Act, FCRA）、反差別法（例えば、Equal Credit
21 Opportunity Act, ECOA、Title VII of the Civil Rights Act of 1964）の適用に関する議論

22 ※2 企業が与信、雇用、保険、ハウジングなどの適格性（eligibility）審査にプロファイリングを用いる際
23 は、FCRA の適用を受け、被評価者に対し告知や異議申立ての機会を与えなければならない可能性がある
24 ことを示唆している。

25 →このような方向性も、プロファイリングの手続を「ホワイトボックス化」する試みの一環と
26 して、あるいは、不条理世界を抑止し、自律的で主体的な生き方を保障するための試みの一
27 環として理解できよう。

30 V. 終わりに

31 32 ●EU とアメリカの動向

33 →AI の発達が現在のペースで進めば、そう遠くない未来、我々の人生の重要な場面で、あ
34 るいは民主主義の重要な場面で、人工知能ベースの予測的アルゴリズムが決定的に重要な
35 役割を果たすことになるろう。このことは決してバラ色の未来を保証するものではない。それ
36 は、AI の確率的な判断によって排除される集団を生むなど、近代立憲主義の基底的価値を

1 揺るがしかねない問題を提起しうるからである。EUやアメリカでは、既にこうした問題に
2 真剣に取り組み始めている。EUが、今後、個人情報・プライバシー保護法制の世界的な準
3 拠点となるであろうGDPRで、「プロファイリング」を定義し、これに特別の規制を加え
4 たのはこの流れの重要な証左である。

5 6 ●日本の現状——「本丸」に関する議論の不足

7 →他方、日本では、改正個人情報保護法のベースとなった2014年6月の制度改正大綱が、
8 プロファイリングへの対応を「継続して検討すべき課題とする」としたものの、改正法に
9 プロファイリングに関する規定が組み込まれることはなかった。かくして現行法制下では、適
10 正に取得された個人データから、いかなるアルゴリズムによって、いかなる個人的側面が
11 プロファイリングされるのか、またこの結果が事業者等の決定と具体的にどう結び付くのか
12 が「ブラックボックス化」された状況にある。学界においても、筆者を含む一部論者がプロ
13 ファイリングの問題点を強調し始めたものの、議論の焦点は、未だ情報漏洩や第三者提供等
14 に置かれているとあってよからう。

15 しかし、そもそも我が国で情報プライバシー権論が盛り上がった際に最も懸念されたの
16 は、高度情報社会において、我々の情報が集積・連結・統合・分析され、それにより我々の
17 個人的側面が知らぬ間に暴かれることだったのではないか(高木浩光)。それは、要するに、
18 プロファイリングである。そして、このことを回避するために、第三者提供を制限するなど、
19 一所懸命、外堀を固めたという側面があったのではないか。そうすると、プロファイリング
20 の規制を欠いた現行法制は、外堀を頑張って固めているが、肝心の本丸の守りが空っぽ、と
21 いう極めて奇怪な——そもそも何のために外堀を固めたのかさえわからないミステリアス
22 な——建付けになっているともいえる。もちろん、従前のプロファイリングは、人間の直観
23 等に依存した部分が多く、予測の精度という点でこれを規制する必要性が現在よりも高く
24 なかったかもしれない。また、その明確な方法も確立していなかったため、規制の仕方も不
25 明瞭であったかもしれない。その意味で、本丸それ自体を守る必要性と、その守り方が判然
26 とせず、とりあえず外堀を固めるしなかったというところがある。

27 しかし、現在は状況が大きく異なる。本報告で指摘したような問題状況を踏まえながら、
28 本丸を本丸として守るための法整備を進めるべきである。

29 30 ●具体的な論点

- 31 ・「個人情報保護法」による規制が妥当か(あるいは業種ごとの“利用”規制か。個情法の
32 射程、あるいは同法とのすみわけをどうするのか。基本法の必要なども含め要検討)
- 33 ・情報取得の規制(取得規制)／プロファイリング規制(実施規制)／結果の利用規制(利
34 用規制)のバランスをどうとるのか
- 35 ・予測精度と公正(憲法価値)とのバランス(特定情報の排除による予測精度の低下をどう
36 評価するのか)

- 1 ・デュー・プロセスの保障
- 2 ・透明性とアカウンタビリティ（ゲーム問題、取引の秘密 etc.）
- 3 ・法とコード（アーキテクチャ）との架橋（「翻訳」→「翻訳家」の育成。法の役割〔価値
- 4 決定〕とアーキテクチャ〔執行〕の役割。新たな立法 - 行政関係論へ？）
- 5