

# AIによる差別と公平性 — 金融分野を題材に

九州大学  
法学研究院  
准教授

成原 慧



桃尾・松尾・難波  
法律事務所  
パートナー弁護士

松尾 剛行



## ～要旨～

本稿では、「AIによる差別」の要因を整理し、「AIによる差別」のリスクに対処するためのソフトロー及びハードローを概観した上で、金融分野においてAIを活用するに当たって留意すべき「AIによる差別」のリスクを明らかにし、「AIによる差別」のリスクへの対処のあり方について検討した。その結果、金融分野における「AIによる差別」のリスクに対処するに当たっては現行法の規制のみでは不十分であり、人間の介入を求める権利の導入など立法論を含めた検討が必要となることを明らかにした。

## 1 はじめに

今日では、AI(人工知能)がチャットボット、与信管理、不正検知など、金融分野を含むさまざまな場面で活用され、業務が効率化されるとともに、利用者に最適化されたサービスが提供されるようになってきている。他方で、AIにより女性やマイノリティの人々に不公平な判断が行われるなど、「AIによる差別」のリスクも顕在化しつつある。そこで、本稿では、「AIによる差別」の要因を整理し、「AIによる差別」に関する原則・指針や法制度を概観した上で<sup>1)</sup>、金融分野においてAIを活用するに当たって留意すべき「AIによる差別」に関する問題を明らかにし、「AIによる差別」のリスクへの対処のあり方を示したい。

## 2 AIによる差別とその要因

### (1) AIによる差別

AIは、人事評価、与信管理、量刑判断などの

場面で活用されるようになってきているが、その際に差別的な判断を行ってしまうおそれがあることが明らかになっている。「AIによる差別」の主な要因としては、①アルゴリズムの設計に起因する差別、②学習するデータに起因する差別、③集団の属性に基づく判断に伴う差別、④人間による責任転嫁をあげることができる。

### (2) アルゴリズムの設計に起因する差別

従来のコンピュータ・プログラムやそれを実装したシステムの場合と同様に、AIも、アルゴリズムの設計により、差別的な判断を行う可能性がある。すなわち、開発者が差別的な意図をもってアルゴリズムを設計したことなどにより、AIが差別的な判断を行うおそれがある。例えば、AIの開発者が特定の人種や民族に対して偏見を持っていて、特定の人種や民族の人々に対して不利になるようにプログラムを書いた結果、そ

のように設計された AI が差別的な判断を行ってしまうといったケースが想定される。もっとも、現実にはこのような悪意を持った開発者は多くないはずであり、現実の問題になることは限られているだろう。他方で、アルゴリズムの設計の際には、意図せずして差別的なアルゴリズムを生み出してしまうという事態も考えられる。例えば、AI を開発する企業のエンジニアが男性だけで構成されている場合には、ともすると女性が置かれている状況に対する視点が欠けてしまいがちで、無意識のうちに女性にとって不利なアルゴリズムの設計をしてしまうという可能性も考えられる。こうしたリスクを防止するためにも、AI の開発者等のダイバーシティを確保・促進することが求められるだろう<sup>2)</sup>。

### (3) 学習するデータに起因する差別

AI は、データからの学習により自らの出力やプログラムを変化させる性質を有している。それゆえ、AI は、データから学習することにより自らの機能を継続的に向上させていくことができる。反面で、このような性質により、AI は、そのアルゴリズムが適切に設計されたとしても、偏ったデータを学習することなどにより、差別的な判断を行ってしまうおそれがある。学習するデータに起因する差別の要因は、さらに、(i) データの代表性の欠如、(ii) データに反映された既存の社会のバイアス、(iii) 相関関係に基づく不正確な予測に分類することができる。

(i) データの代表性の欠如は、現実の社会を適切に反映していないデータセットから AI が学習した結果、AI が差別的な判断を行ってしまうという問題である。つまり、AI の学習に用いられるデータの中にさまざまな人々のデータが公正に代表されていないと、AI が差別的な判断を行ってしまうおそれがあるのだ。例

えば、スマートフォンのアプリを通じて住民から調査を行う場合、スマートフォンの保有率の低い低所得者や高齢者のニーズが反映されにくくなってしまうおそれがある。AI にデータを学習させる際には、このようにデジタル・デバイドに起因するデータの偏りにも注意する必要がある<sup>3)</sup>。また、顔画像を認識する AI を開発する際に、白人の顔画像のデータを中心に学習させると、黒人やアジア系の人の顔を正確に認識しにくくなってしまっておそれもある。

(ii) 現実の社会を適切に代表するデータであっても、現実の社会自体にバイアスがある場合には、それを反映して AI が差別的な判断を行ってしまうおそれがある。つまり、データの元になっている社会の構造自体に不公平なバイアスがあると、そうした従来の社会の差別的な構造をデータが反映し、それを学習した AI が差別を再生産してしまうおそれがある。例えば、Amazon が、自社の従業員採用において過去の応募者の履歴書のデータを元に応募者の履歴書を評価する AI を開発しようとしたところ、過去の応募者に男性が多数を占めていたため、AI が男性の応募者を優遇し、女性の応募者を不利に扱うおそれがあることが明らかになったため、開発を中止したと報じられている<sup>4)</sup>。また、原因は特定されていないものの、AI により算定された女性のクレジットカードの利用限度額が男性の利用限度額よりも低くなる傾向があるといった問題も生じている<sup>5)</sup>。

(iii) 学習するデータに起因する差別の要因として、不正確な予測による差別があげられることもある。AI によるデータ分析は、相関関係を明らかにすることができるが、データ分析により明らかにされた相関関係は、必ずしも

意味のあるものとは限らず、因果関係を伴っているとも限らない。例えば、地域ごとの新型コロナウイルス感染症に関するキーワードの検索リクエスト数は、ある時期には、地域ごとの感染者数と相関することもあるかもしれないが、状況が変化すると、相関しなくなるかもしれない。それにもかかわらず、両者に相関関係があるとの予断の元に予測を行うと不正確な結果を招くおそれがある。こうした不正確な予測に基づいて政府や企業が個人の権利利益に関する判断を行うことになれば、一定の地域に居住する人々は感染リスクが高いといった不正確な判断を行うことなどにより、一定の属性を持つ人々に差別的な影響が及ぶおそれもある<sup>6)</sup>。

#### (4) 集団の属性に基づく判断に伴う差別

「AIによる差別」を引き起こすより根本的な要因として、個人の属する集団の属性に基づく判断が問題とされることがある。すなわち、AIは、個人をさまざまな属性をもったデータの束として認識するため、AIは、個人をその人の有するさまざまな属性に基づいて確率的に判断することになる。このように、AIは、個人を属性の束として扱うことにより、個人のかげがえのなさを捨象してしまうというのである。AIは、同じ属性を分かち持つ集団の一般的傾向により、個人の能力や性質を抽象的・概括的に判断する傾向がある。このようなAIの判断手法は、身分など集団の属性により個人の能力や性質を抽象的・概括的に判断し、人々の生き方を予め規定してきた前近代社会の評価手法に類似しているとされる<sup>7)</sup>。

#### (5) 人間による責任転嫁

「AIによる差別」の中には、実際には人間が

差別的な判断を行ったにもかかわらず、AIに責任転嫁するケースも考えられる。例えば、実際にはAIの利用者が差別的な判断を行ったにもかかわらず、AIがそのような判断を行ったなどと主張したり、自らがそのような判断をした根拠はAIによるデータ分析に基づいているなどと主張して、自らの責任を回避したり、判断の客観性・科学性を標榜する事態も考えられる。

### 3 AIによる差別に関する原則・指針と法制度

#### (1) AIによる差別の防止を求める原則・指針

近年では国際的にAIに関する原則・指針が策定されるようになってきているが、以上のような問題を念頭に、主要なAI関係の原則・指針においてもAIの判断の公平性の確保や差別の防止が求められている。例えば、総務省AIネットワーク社会推進会議が2019年に取りまとめた「AI利活用ガイドライン」では、⑧公平性の原則が掲げられ、AIの利用者等に「AIシステム又はAIサービスの判断によって個人及び集団が不当に差別されないよう配慮する」ことが求められている。その解説では、「AIの学習等に用いられるデータの代表性やデータに内在する社会的なバイアス等に留意」するとともに、「学習アルゴリズムによるバイアス」に留意することが求められている<sup>8)</sup>。AIの開発者・利用者は、「AIによる差別」を防止するために、こうしたAIに関する原則・指針の内容を参照することが期待される。

#### (2) 法の下での平等と私人間における差別

日本国憲法14条1項は、「すべて国民は、法の下に平等であつて、人種、信条、性別、社会的身分又は門地により、政治的、経済的又は社会的関係において、差別されない」と定めている。憲法の人権規定は一義的に国又は公共団体の統治行動との関係で個人の自由と平等を保障した

ものと解されている。したがって、国や地方公共団体が、規制行政や給付行政においてAIを活用した結果、特定の個人を人種や性別等により差別的に取り扱った場合には、憲法14条1項の定める平等原則に違反する可能性がある。他方で、憲法の人権規定は、国又は公共団体と個人の関係を規律するものであり、民間企業と個人など私人相互の関係には直接適用されないと解されてきたため<sup>9)</sup>、企業など私人が「AIによる差別」を引き起こしたとしても、直ちに憲法の平等原則に違反するとは言い難いだろう。もっとも、企業など私人がAIを利用する場合であっても、民法の一般条項などを介して憲法の平等原則の趣旨が間接的に適用され、「AIによる差別」が違法とされる余地もある<sup>10)</sup>。また、信条等による差別や男女差別を禁じる労働法や各種の業法の規制<sup>11)</sup>に基づき、企業が「AIによる差別」を防止することが法的に要請される場合も考えられる<sup>12)</sup>。

### (3) 個人情報保護法による差別の防止

AIは個人情報を含むデータを利用することから、「AIによる差別」を防止するに当たっては、個人情報保護法の役割も期待される。個人情報保護法1条は、「個人情報の有用性に配慮しつつ、個人の権利利益を保護すること」を同法の目的と定めている。「個人の権利利益」の中には、プライバシーに加え、不当な差別を受けないことなども含まれると解されている。すなわち、個人情報保護法は、不当な差別など個人の権利利益の侵害を予防することを目的として、個人情報の取扱いに関する義務の遵守を求めているといえる<sup>13)</sup>。

個人に対する不当な差別や偏見を予防する見地から、平成27年改正により「本人の人種、信条、社会的身分、病歴、犯罪の経歴、犯罪により害を被った事実その他本人に対する不当な差別、

偏見その他の不利益が生じないようにその取扱いに特に配慮を要するものとして政令で定める記述等が含まれる個人情報」である要配慮個人情報に関する規定が導入された。要配慮個人情報(同法2条3項)の取扱いについては、取得に原則として本人の同意が求められる(同法20条2項)など、一般の個人情報の場合よりも厳格な規制が課せられている。

また、令和2年改正により、AIを用いたプロフィールによる差別のリスクなども念頭に置いて、不適正な利用の禁止規定(同法19条)が導入されている<sup>14)</sup>。

### (4) 特定デジタルプラットフォーム透明化法による公正性の確保のためのアルゴリズム開示

特定デジタルプラットフォームの透明性及び公正性の向上に関する法律5条2項1号ハは、「検索により求める商品等に係る情報その他の商品等に係る情報に順位を付して表示する場合における、当該順位を決定するために用いられる主要な事項(商品等提供利用者からの当該特定デジタルプラットフォーム提供者に対する広告宣伝の費用その他の金銭の支払が、当該順位に影響を及ぼす可能性がある場合には、その旨を含む。)」の開示を求めている。すなわち、検索結果等における商品掲載順序の決定に用いられる主要な基準(広告宣伝費用などの支払が掲載順位に影響する場合には、その旨)を一般利用者に開示すべきとされている。同法のアルゴリズムの重要部分の開示義務はあくまでも、特定の文脈における開示義務に過ぎないものの、融資判断等においても、何が重要な判断要素とされているか等を開示されることが求められる可能性がある<sup>15)</sup>。

#### (5) 独占禁止法による（競争の）公正性の確保

あくまでも競争の公正性という観点に基づくものであるが、アルゴリズムを利用して競争の公正性を害する行為を行った場合、例えば差別的取り扱いの禁止や優越的地位の濫用として違法となる可能性がある。食べログ事件判決（令和4年6月16日）では、飲食店ポータルサイトのアルゴリズム変更が優越的地位の濫用に当たるとされた。

#### (6) 欧米の立法動向

欧米では「AIによる差別」に対応するための立法に向けた検討も進められている。例えば、欧州委員会は、2021年にAIに関する新たな法的枠組みを構築するための規則案を提案した。規則案では、AIシステムにより差別的な判断が行われるリスクがあるといった認識を踏まえ、算出される社会的スコアにより特定の個人又は集団に不当に不利益が及ぶおそれがある場合などに公的機関が個人の信頼性を評価・分類するためにAIシステムを提供・利用することなどが禁じられている。また、規則案では、自然人の遠隔生体識別、信用評価やクレジットスコア算出などのために利用されるAIシステムを、高リスクのAIシステムと位置づけ、高リスクのAIシステムについて、ライフサイクル全体におけるリスク管理システムの構築・実装、開発時のデータガバナンスならびに情報の記録・開示及び人間による監督を可能とする設計・開発などの義務を課している<sup>16)</sup>。

米国ホワイトハウスは、2022年10月に、「AI権利章典のための素案」を公表している。素案では、雇用や与信に用いられるアルゴリズムが、既存の差別を反映・再生産し、新たなバイアスと差別を埋め込んでいるという認識などを踏まえ、5原則の一つとして「アルゴリズムによる差別

からの保護」が掲げられている。同原則では、AI等を活用した自動化システムの設計者、開発者及び実装者に対し、アルゴリズムの公平性評価など、個人及び共同体をアルゴリズムによる差別から保護するための予防的かつ連続的な措置を取るとともに、当該システムを公平な方法で利用・設計することを求めている<sup>17)</sup>。

## 4 金融機関の利用するAIによる差別と公平性

前章で説明した一般的なAIによる差別と公平性の問題を、以下では具体的な金融機関実務の場面に落とし込んで検討したい。具体的には、融資判断及び不正検知の2つの場面を検討する。

### (1) 融資判断において利用されるAIによる差別と公平性

これをAIによる融資判断と呼ぶかはAIの定義の問題であるが、従前から融資判断にスコアリングが取り入れられたところ、スコアリングに過度に依存した融資判断の誤りが批判されてきたところである<sup>18)</sup>。

その後、融資判断に関する情報技術の高度化<sup>19)</sup>により、従前批判された問題点の解決が期待されるものの、例えば、一部の属性を有する融資申込者を「高リスク」と分類して一律排除した結果として、従前より大幅に返済率が上がったという状況を「解決」として高く評価すべきかはさらに検討を要する。すなわち、そのような申込者がAIによる差別の被害を受けている可能性がある。

融資判断において利用されるAIによる差別としては、人種、国籍、性別等による融資拒否が典型的なものである。そして、現在の社会構造（例えば、米国において人種ごとによりはっきりとした貧富の差が存在すること）の反映とし

て、それらの属性を理由に「ハイリスク」と分類して融資をしないことは、返済率の向上という目的に資する面があることは指摘せざるを得ない。もっとも、そのような属性による融資判断は、上記2及び3で述べたような差別や不公平な扱いにつながり得る。以下では、AIの設計には注意を払う(2(2)の設計による差別の問題は生じない)ことを前提として具体的に検討していこう。

AIを融資の際の判断に用いると、2(3)で見た学習するデータに起因する差別が生じうる。学習するデータに起因する差別のうち、(i)データの代表性の欠如に起因する差別については、例えば、クレジットカード利用履歴を元と与信判断を行う場合、特定の属性を有する個人がクレジットカードを発行してもらいにくくなり、仮に発行されても限度額が低いため、良好なクレジットカード利用履歴が得られなくなった場合に、AIが学習ないし解析の対象とするデータプールと、現実世界における各コミュニティのデータとの関係が歪み、差別や誤りが容易に生じやすくなるという問題が想定される<sup>20)</sup>。(ii)データに反映された既存の社会のバイアスに起因する差別については、従来特定の属性を有する人々に対して不当に不利益を与えるように人間(銀行員)が融資判断をしていた場合において、当該融資判断の内容をAIに学習させることで当該差別が再生産されるといったことが懸念される。(iii)相関関係に基づく不正確な予測に起因する差別については、特定の属性を持つ人々の融資返済率が低いと、それと当該属性の間に何ら因果関係はないにもかかわらず、AIがうわべだけの相関関係を読み取ってしまって、誤った融資拒絶が発生する可能性がある<sup>21)</sup>。この点につき、米国FTCの報告書が、うわべだけの相関関係に基づいて不正確な判断が行われるリスクに

注意する必要があるとした上で、健康、与信、雇用などに関わる重要な決定にアルゴリズムとデータが用いられる場合には、アルゴリズムとデータに対して人間が監督を行うよう推奨していることにも留意が必要である<sup>22)</sup>。

加えて、2(4)で見た集団の属性(セグメント)に基づく判断に伴う差別も生じうる。山本龍彦は、このような確率的評価のみで個人の(与信判断における信用を含む)信用力が判断されるとなると、その人と属性を共有する者の一般的傾向によって、個人の能力が概括的・抽象的に判断され、それによって個人の尊厳が害される恐れがあるとした上で、人間の介入を求める権利を認めるGDPRの規定を「最終的には人間が、被評価者一人ひとりと向き合い、彼らが発する肉声に耳を傾けることを法律上要求したものと考えることができる。」と評する<sup>23)</sup>。融資判断を完全に自動化し、セグメントのみに基づき判断し、人間に介入して欲しいという異議を一切受け入れなければ、私人間にも適用される個人の尊厳(憲法13条)の価値に鑑みて、法的な問題が生じ得るだろう<sup>24)</sup>。この点に関し、山本は被評価者(例えば融資申込者)として、AIの予測や評価に実質的に反論を加えられるようにするため、アルゴリズムのロジックの重要部分が開示されることの重要性を論じる<sup>25)</sup>。なお、ロジックが開示され、透明化することは、本当は人間が差別したにもかかわらず、それをAIのせいにする(責任転嫁)ことの回避にもつながる<sup>26)</sup>。

なお、融資AIが特定のAIベンダによる寡占状態となり、ほとんどの金融機関が同一のAIエンジンを利用するようになれば、「属性(例えば人種、国籍、性別等)について特定の組み合わせを有する人は、どの銀行に行っても融資を断られ続ける」という状況(バーチャルスラム<sup>27)</sup>)が生じ得ることも重要な懸念点である。

## (2) 不正検知において利用される AI による差別と公平性

不正検知や監査業務においても、AI が活用されている<sup>28)</sup>。金融機関が日常的に取り扱う大量の取引の全てを人間の目でチェックして不正を検出するのは労力がかかり過ぎる。ビッグデータに基づき AI アルゴリズムが統計的にリスクが高い取引を選び出すことで、人間がレビューすべき対象を選別すれば、効率的に不正を検知し、監査を行うことができる<sup>29)</sup>。金融機関においては、例えば、アンチマネーロンダリング<sup>30)</sup>やインサイダー取引防止<sup>31)</sup>等の観点で AI が利用されている。このうち、上記の融資の議論が転用できる限りにおいてはあえて繰り返しを避けるために議論を省略し、不正検知特有の問題にフォーカスして分析をしたい。

まず、不正検知の場合には、その結果により懲戒や刑罰等の厳しい結果が生じ得ることから、冤罪の被害が重大であることを指摘することができる。オランダでは、社会保障や税の不正受給・還付等を特定するための機械学習アルゴリズムが、貧しい地域に住む人に対し不当に育児手当を返還させる等の問題を起こし、大きな問題となった<sup>32)</sup>。日系企業が買収した英国のコンピューター企業が提供した郵便局システムの欠陥により 700 人以上の郵便局長らが横領や不正経理の無実の罪を着せられたとされる事案もある<sup>33)</sup>。

また、横領等の不正を理由に刑事罰を科す場面や、従業員の不正を理由に懲戒解雇をする場面等を想定すると、その際に要求される説明において、単に「AI が不正だと述べたから」では不十分と言わざるを得ない。つまり、重大な結果を招来する場面が想定される不正検知等においては、刑事訴訟法や労働契約法等に基づき法的に説明が要求される所、AI、特に機械学習型の AI の説明だけでは不十分な可能性がある<sup>34)</sup>。

AI を単なる不正の兆候を探知する契機や大量の資料の選別の支援ツールとして利用し、実際の説明に利用する資料は人間が選別して提示することが現実的であろう。

最後に、プライバシーへの制約に関して、①日常的な調査と不正の兆候の検知後の調査の相違と、② AI によるモニタリングと人間によるモニタリングのプライバシー制約の程度の問題について触れたい。①日常的な調査においては、特に不正の兆候がない中のモニタリングであり、一般には必要性、方法と程度の相当性、根拠等が重要である<sup>35)</sup>。これに対し、有事対応としての調査では証拠収集の必要性が高く、もちろん方法と程度の相当性、根拠等も問われるものの、相対的には正当とされやすい<sup>36)</sup>。② AI によるモニタリングは網羅的でプライバシーへの制約の程度が高いという考えと、感情を持たない AI によるモニタリングはプライバシーへの制約の程度が低いという双方の考えがあるが、このような一般論での議論は無意味であろう。むしろ人間と AI の協働を前提に、その具体的協働がどのような態様かを考えるべきである。例えば、人間だけでメールをモニタリングすれば、無関係なメールを見る可能性も高いが、AI が選別したメールを人間が見ることとすれば、効率性が向上するだけではなく、監視の必要性が高いメールのみを見ることになるので、プライバシーの観点からも適切とされることが増えるだろう<sup>37)</sup>。

## 5 総括

「AI による差別」や AI による公平性の障害は既に問題となっており、金融の場面においても問題となり得る状況が存在する。その対応策として、現行法として個人情報保護法 19 条の規律、特定デジタルプラットフォーム透明化法及び独占禁止法等が存在するが、例えば、個人情報

を利用しないアルゴリズムに対して個人情報保護法19条は無力であるし、特定デジタルプラットフォーム透明化法の適用範囲は特定デジタルプラットフォーム提供者に限られており、また、競争の公正が問われる場面以外の場面に独占禁止法は働きにくい。このように、現行法では「AIによる差別」への対応として必ずしも十分ではない。

そこで立法論として、「AIによる差別」のリスクに対応するために、欧州AI規則案同様に、許容し得ないリスクを有するAIを禁止するという選択肢も考えられるものの、禁止すべきAIの範囲についてはAI技術の健全な発展の観点を踏まえ、慎重に検討すべきであろう。むしろ、多くのAIは一律に禁止するよりはむしろ、人間とAIがそれぞれの欠点を補い合うケンタウロスモデルを採用することにより差別や不公平な判断を防止することが合理的であろう。このような見地からは、GDPRの定める人間の介入を求める権利に相当する規定を日本法に導入することが考えられる。かかる規定については、個人情報保護法に導入するという選択肢とともに、例えば融資について人間の介入を求めることが重要であれば、銀行法等の業法に導入するという選択肢も考えられる。

#### 【注】

- 1) 本稿の第2節・第3節の一部は、成原慧「『AIによる差別』にいかに向き合うか」宮下萌(編)『テクノロジーと差別』(解放出版社、2022年)を元にしている。
- 2) European Commission (High-Level Expert Group on AI), Ethics guidelines for trustworthy AI 23 (2019); Sarah Myers West, Meredith Whittaker & Kate Crawford, Discriminating Systems: Gender, Race and Power in AI pp.15-19 (2019).
- 3) FTC, Big data: A Tool for Inclusion or Exclusion? pp.27-28 (2016).
- 4) Jeffrey Dastin, Amazon scraps secret AI recruiting tool that showed bias against women, Reuter, Oct. 11, 2018.
- 5) Will Knight, The Apple Card Didn't 'See' Gender—and That's the Problem, WIRED (Nov. 19, 2019).
- 6) FTC, *supra* note 3, at 29-31.
- 7) 山本龍彦「ロボット・AIは人間の尊厳を奪うか?」弥永真生=宍戸常寿(編)『ロボット・AIと法—ロボット・AI時代の法はどうなる』(有斐閣、2018年) pp.86-88 以下参照。
- 8) 総務省AIネットワーク社会推進会議「AI活用ガイドライン」pp.23-24 (2019年) 参照。
- 9) 最大判昭和48年12月12日民集27巻11号1536頁(三菱樹脂事件)。
- 10) 日産自動車事件(最判昭和56年3月24日民集35巻2号300頁)において最高裁は、女性従業員の定年年齢を男性従業員より低く定めていた就業規則について、「性別のみによる不合理な差別を定めたものとして民法90条の規定により無効であると解するのが相当である(憲法14条1項、民法1条ノ2参照)」と判示している。
- 11) 例えば、労働基準法3条、男女雇用機会均等法5条・6条、障害者差別解消法8条、電気通信事業法6条、金融庁所管事業分野における障害を理由とする差別の解消の推進に関する対応指針等参照。
- 12) 例えば、企業の人事考課において用いられたAIの評価や判断に差別や恣意性があった場合には、人事権の濫用として違法となりうる。松尾剛行『AI・HR テック対応 人事労務情報管理の法律実務』pp.230-239 (弘文堂、2019年)、「労働政策審議会労働政策基本部会報告書～働く人がAI等の新技術を主体的に活かし、豊かな将来



- を実現するために～」(https://www.mhlw.go.jp/stf/newpage\_06697.html) 及び IBM 事件 (http://www.jmitu-ibm.org/2022/04/8461.html、http://www.jmitu-ibm.org/2020/04/7687.html) 参照。
- 13) 富安泰一郎＝中田響 (編著)『一問一答 令和3年改正個人情報保護法』p.98(商事法務、2021年)。
- 14) 個人情報保護委員会「個人情報の保護に関する法律についてのガイドライン(通則編)」p.39(平成28年11月(令和4年9月一部改正))も参照。
- 15) 筆者(松尾)がいわゆる食ベログ事件判決(後掲3(5)参照。)を獲得した際も、食ベログ側からアルゴリズムに関する一定の情報の開示を引き出したものの、必ずしも当然に開示義務を負っているわけではないことが前提であったことから、様々な工夫を要した。
- 16) *Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Com (2021) 206 Final (Apr. 21, 2021)* .
- 17) The Whitehouse, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* (Oct. 2022) .
- 18) 例えば「同行は、過大な事業規模の追求、スコアリングモデルのみに依存した融資審査・管理等に起因して大幅な損失を計上してきた。」とする金融庁「株式会社新銀行東京に対する行政処分について」2020年12月26日(https://www.fsa.go.jp/news/20/ginkou/20081226-5/01.pdf) や、2000年代前半のスコアリングモデルが失敗した原因を「スコアリングモデルが未成熟」「定性分析を行っていない」「粉飾決算の可能性」「母集団の違い」等に分けて論じるりそな銀行「AIを活用した信用評価手法の現状とこれから」日本銀行 金融高度化センターワークショップ「AIを活用した金融の高度化に関するワークショップ」第3回2019年2月18日(https://www.boj.or.jp/announcements/release\_2019/data/rel190215d2.pdf) p.10を参照。
- 19) 「銀行が、クレジット・スコアリング事業者と協働して、決済情報のビッグデータとAIを組み合わせると信費用(信用コスト)の発生確率の分析精度を向上させることが検討されている。こうしたことが実現すれば、これまで融資実行が困難と判断された利用者への資金供給を行うことにより、(ii)利用者の利便の向上や(iii)金融サービスの高度化、ひいては経済全体の発展に資することが期待されるほか、(i)銀行業の効率性・収益性の向上も期待されよう。」とする「金融サービスにおける顧客情報の利用を巡る法律問題研究会」報告書:顧客情報の利活用に関する行為規範のあり方(https://www.imes.boj.or.jp/research/papers/japanese/kk40-1-1.pdf) p.24及び、金融庁「『AIとデータ活用の融合によるフローデータを用いたオンラインレンディングの進展』多様なフィンテックステークホルダーとの対話から見えた10の主要な発見(Key Findings)」(https://www.fsa.go.jp/news/r1/sonota/FIH\_Report.pdf) p.22.
- 20) 山本龍彦『AIと憲法』p.44。(日本経済新聞出版、2018年)も参照。
- 21) なお、山本・前掲注(7)96頁は、「融資判断の際に使われるAIの信用力評価に、“特定の相手にどれぐらいの頻度で連絡しているか”という通信履歴が秘密裡に使われていたとしよう。このような定期的な連絡をしないA氏は、知らぬ間に信用力スコアを下げられることになる。」とする。
- 22) FTC, *supra* note 3, at 32.
- 23) 山本・前掲注(7) p.89参照。
- 24) なお、AIの誤った判断によりアカウント停止等を行い、人間の介入を求める異議を聞かない国際動画共有サイトを筆者(松尾)がVtuberの代理人として提訴した事案につき、松尾剛行「ブラッ

トフォーム事業者によるアカウント凍結等に対する私法上の救済について」情報法制研究 10号 p.66 参照。

25) 山本・前掲注 (7) p.91 参照。

26) この点は、上記 3 (4) 及び山本・前掲注 (7) p.91 注 28 も参照。

27) 同上 p.91 参照。

28) 公認会計士・監査審査会「令和 4 事務年度監査事務所等モニタリング基本計画」2022 年 7 月 15 日 ([https://www.fsa.go.jp/cpaaob/shinsakensa/kihonkeikaku/20220715/2022\\_kihonkeikaku.pdf](https://www.fsa.go.jp/cpaaob/shinsakensa/kihonkeikaku/20220715/2022_kihonkeikaku.pdf)) p.1 参照。

29) ただしゲーミングと言って、アルゴリズムを逆手にとってリスクが低そうに見せかけるリスクに留意が必要である。

30) 一般社団法人全国銀行協会「AML / CFT 業務の高度化・共同化に係る新会社の設立について」令和 4 年 10 月 13 日 (<https://www.zenginkyo.or.jp/news/2022/n101302/>) 参照。

31) 株式会社 SBI 証券・日本電気株式会社「SBI 証券、NEC と共同で、国内で初めてインサイダー取引の審査業務に AI を導入 ～不公正取引度合いのスコアリングで一次審査時間を約 90% 短縮～」2021 年 9 月 14 日 ([https://jpn.nec.com/press/202109/20210914\\_02.html](https://jpn.nec.com/press/202109/20210914_02.html))

32) 有限責任あずさ監査法人「AI の適切性検証への取組み」2022 年 2 月 15 日 ([https://www.soumu.go.jp/main\\_content/000826719.pdf](https://www.soumu.go.jp/main_content/000826719.pdf)) p.5 及び岩佐淳士「突然、詐欺犯のぬれ衣…オランダ、AI が標的にした 2 万 6,000 人」2022 年 6 月 1 日 毎日新聞 (<https://mainichi.jp/articles/20220530/k00/00m/030/196000c>) 参照。

33) Mariko Oi, *How a Japanese firm became part of the Post Office scandal*, BBC October 14, 2022 (<https://www.bbc.com/news/business-61020075>)

34) 松尾・前掲注 (12) p.39 参照。

35) 同上 pp.187-188 参照。

36) 同上 pp.303-304 参照。

37) 同上 pp.215-216 参照。

---

なりはら さとし  
九州大学法学研究院准教授。  
東京大学大学院情報学環助教、総務省情報通信政策研究所主任研究官等を経て、2018 年より現職。

【専門】  
情報法。

【主な著書】

『表現の自由とアーキテクチャ』勁草書房、2016 年  
『AI がつなげる社会— AI ネットワーク時代の法・政策』(共著) 弘文堂、2017 年  
『AI で変わる法と社会』(共著) 岩波書店、2020 年  
『テクノロジーと差別』(共著) 解放出版社、2022 年など

まつお たかゆき

桃尾・松尾・難波法律事務所パートナー弁護士・NY 州弁護士、慶應義塾大学特任准教授、法学博士。

【主な取扱案件】

食ベログ事件 (東京地判令和 4 年 6 月 16 日)。

【主な著書】

『紛争解決のためのシステム開発法務』(共著) 法律文化社、2022 年  
『AI・HR テック対応 人事労務情報管理の法律実務』弘文堂、2019 年  
『金融機関における個人情報保護の実務』経済法令研究会、2016 年 など

---