

人工知能による選別と翻弄される労働者

——法は何をすべきか？——

竹 地 潔

富山大学紀要. 富大経済論集 第65巻第2号抜刷（2019年12月）

富山大学経済学部

人工知能による選別と翻弄される労働者

——法は何をすべきか？——

竹 地 潔

キーワード：アルゴリズム，一般データ保護規則，AI面接，HRテック，機械学習，差別，人工知能，ディープラーニング，データ保護影響評価，バイアス

I 問題の所在

現在、企業の採用活動における「AI面接」の活用が注目を浴びているように、人事労務管理業務全般に人工知能による人事データ分析が利活用され始めている。人工知能による人事データ分析のうたい文句は、人間の管理監督者等による「主観的」な評価に左右されることなく、「科学的」な分析を通じて労働者に対し「客観的」な評価を下すことができる、というものである。

しかし、アルゴリズムの開発段階で、バイアス（偏り，偏見）が意図せずに入り込んだり，読み込んだデータの種類によっては人間の予測しない誤った方向に学習が進んだりすることが原因で，人工知能によるデータ分析が特定の属性を有するグループに対し不利益ないし差別的な判断を下すことがある，との事例が報告されている¹⁾。たとえば，最近の代表的な事例として，アマゾン社のケースがある。人工知能を用いて優秀な人材を探し出すシステムの開発において，技術関連の職種について過去十年分の提出された履歴書のパターンをコンピュータモデルに学習させたところ，当該職種への応募者がほとんど男性であったため，当該システムが男性を採用することが好ましいとの結論を下し，性中立的な判断ができないことが明らかとなった。そのため，同社は，プログ

ラムを修正したものの、別の差別を生じさせる恐れがあるとし、当該システムの開発を断念した²⁾。

このように、人工知能による人事データ分析は、必ずしも客観的に正確かつ公正な評価を下すことができるとはいえない。そのため、当該データ分析は、バイアスのかかった評価を行い、労働者の権利や自由および正当な利益に重大な影響を及ぼす恐れがある。したがって、労働者の権利や自由の保護の観点から、人工知能による人事データ分析の利活用およびそれに基づく人事労務管理上の決定に対し何らかの法的規制を加える必要があるとの見解が強まっている。

以下、本論は、まず、人工知能による人事データ分析の利活用とその普及を概観し、人工知能による分析技術とその問題点を指摘する。次に、わが国の法的対応の現状を確認したうえで、先進的な取り組みとしてのEU一般データ保護規則を学び、それに対し批判的検討を加える。最後に、EUの先進的な取り組みを踏まえ、わが国の法的規制についての基本的な方向性を示すこととする。

II 人工知能による人事データ分析の利活用とその普及

約20年前から米国では、「HR テック (HR Tech)」つまり、情報技術 (IT) を使って人事評価や人材の採用・育成等の人事労務管理業務全般を改善するサービスが利活用され始め、現在世界的に大きな注目を集めている。HR テックとは、人事 (Human Resources = HR) とテクノロジー (Technology) を組み合わせた造語で、人工知能等の最先端テクノロジーを使って人事労務管理業務全般の効率化とクオリティーの向上を図るための手法である。たとえば、人工知能による人事データ分析に基づき、将来活躍しそうな有能な人材を発掘したり、また、離職の可能性があるような従業員を検出し、事前に離職防止策を打ち出したりするといった将来予測などを行うものである³⁾。

近年わが国も、米国に比べて大幅に後れをとりつつも、人事労務管理業務全

般に「HR テック」が導入され、人工知能による人事データ分析が利用され始めている。

まず、採用業務について、従来から、エントリーシートの確認や面接日程の調整などに手間がかかることや、応募者の評価結果について、異なる面接官によってバラツキが生じることから、業務の効率化と合否判定の精度の向上が課題となっていた。これらの課題を解決するために、人工知能による分析技術を活用した「定量的」かつ「客観的」な採用選考手法が導入され始めている。たとえば、人工知能によって、エントリーシートの記載内容を評価し選別する手法に始まり、応募者の入社後の活躍度を予測する手法、および、個人と人間関係・社風等とのマッチング度合いを事前に測定する手法などである。これらの手法の利活用によって、エントリーシートの確認に要する時間の大幅な短縮化、応募者との出会いの増加と新たな人材発掘、および、企業（または職務）と人材とのミスマッチの防止につながる効果が大きいと期待されている⁴⁾。

次に、従業員の配置・異動についても、従来から、人事担当者の「勘」や「経験」で決定されがちであったが、多様な職種やポストにより適切な人員配置を行うために、人材のさまざまな資質が見える化し、職務の特性と掛け合わせて分析することにより、「人」と「職務」のフィット（Job Fit）の程度を定量的に測定する手法が活用され始めている。この手法で得られた情報は、人員配置に関してより精度の高い判断を行うことを可能にするとともに、個々の従業員によるキャリアプランニングの検討にとっても大いに参考になりうる、といわれている⁵⁾。

また、近年「健康経営」が提唱される中、ビッグデータ分析エンジンを用いて、従業員の勤怠データや健康診断結果を分析し、メンタルヘルスの予兆を検出し、また、将来の健康状態のシミュレーションを行うことによって、メンタルヘルス不調者への早期対応や、従業員自らによる生活習慣の改善および健康維持・増進への支援を行う取り組みも見られる。さらに、現在の人手不足の中、従業員の定着が大きな経営課題となっていることもあり、突然の退職を回避するた

め、退職しそうな気配のある者を早期に察知し、適切な対応をとるため、人工知能による分析に基づき、勤怠データから4ヶ月後の退職確率を予測する手法も利用されるようになってきた⁶⁾。

以上のように、人事労務管理業務全般において、人工知能による人事データ分析が利活用され、大きな成果が上がることが期待されている。

Ⅲ 人工知能による分析技術と問題点

人工知能による分析で用いられているのは、「機械学習」、つまり、データから規則性や判断基準を学習し、それに基づき未知のものを予測、判断する技術である。以下、採用選考時などにおけるその利用の事例に即して、その内容を説明し、問題点を指摘する。

機械学習の1つである「教師あり学習」の分析手法では、応募者からの採用候補者の選抜に際して、いかなる特徴を優先的に評価すべきかを、開発者が指定しアルゴリズムを作成する。このアルゴリズムによって、所定の学歴や職歴もしくは資格や技能、または、必要とされる適性等を有する「最適」な候補者が選び出されることになっている。

とはいえ、当該アルゴリズムは、人間である開発者が作成したものであり、その作成過程（たとえば、優先的に評価すべき諸特徴の選択）において開発者の先入観、誤解、バイアスが無意識にアルゴリズムに入り込む余地がある。その結果、そのようなアルゴリズムに基づき、特定のグループ（人種、性別、年齢、障害等々）の応募者が選別され、候補者から除外される恐れがある。このように、アルゴリズムは、人間を介して社会における既存のバイアス（偏り、偏見）が反映されうる、と指摘されている⁷⁾。

他方、「教師なし学習」の分析手法では、アルゴリズムが自ら優先すべきデータを見極める。つまり、アルゴリズムが、これまで雇用してきた優秀な従業員の各種のデータから、今後採用する従業員に必要とされる特徴や技能等を導出

する。そして、それらに該当する応募者が「最適」な候補者として選抜される。

しかし、この分析手法については、アルゴリズムに与えられる学習用データが不十分であったり、多様性に欠けていたりすると、バイアス（偏り、偏見）を帯びた解が導き出されうる。つまり、与えられるデータセットの中に、特定の特徴や技能が過小にしか含まれないと、その特徴や技能を軽視する結果に偏る一方、特定の特徴や技能が過多に含まれると、それらを重視する結果に偏ることがある。また、データセットの中に、特定のグループの人々に関するデータが全く含まれていないか、または、わずかししか含まれていない場合は、そのグループの人々が評価対象から外されたり、適正に評価されなかったりすることがある。このように、アルゴリズムへの学習データの与え方によって、アルゴリズムにバイアス（偏り、偏見）が持ち込まれるのである⁸⁾。

さらに、近年脚光を浴びている分析手法として、ディープラーニング（深層学習）がある。それは、脳の神経回路の構造を模倣したニューラルネットワークを用いた機械学習の手法の1つであり、コンピュータが膨大なデータを学習し物事の特徴やパターンを自動的に抽出し、特定の判断や予測に役立てることができる技術である。採用選考、人事評価および配属など人事労務管理業務全般にその利活用が期待されている。

しかし、ディープラーニング技術は、読み込ませたデータの種類によって、人間がまったく予測していない方向へ学習が進む恐れがある。たとえば、偏見や差別、妬みや恨みなどのネガティブな感情を含むデータを読み込まれ続けると、学習にバイアスがかかり、独善的で排他的な思考法をひそかに身につけることもある⁹⁾。

また、ディープラーニング技術は、その特性として、自ら導き出した判断や予測の結果に至る過程について人間の言葉で説明することができず、なぜそのような結果になったのかについて、人間は理解することができない。このことはまさに「ブラックボックス化」の問題である。つまり、同技術で導出された結果のみに基づき雇用上の決定がなされると、労働者は不利益な取扱いを受け、

その理由を知りたいと思っても、その理由を知ることができない¹⁰⁾。(たとえ、労働者がその理由を使用者に尋ねたとしても、ただ、「人間よりも『優れた公正中立』な人工知能による『客観的』な評価に基づくものです」との返事が使用者から来るだけかもしれない。人工知能による判断が、あたかも神々の審判のように。)

他方、雇用上の重要な決定に際して、同技術で導き出された結果をそのまま、その理由を知らずに、鵜呑みにすることは、使用者にとってもリスクがある。そのため、その結果の合理性を人間自らが二重にチェックしなければならず、二度手間になる場合もある。つまり、同技術のメリットである業務の効率化が失われることになる。そのせいで、近年、米企業では、このようなブラックボックス化の解消のため、透明性のより高いアルゴリズムの開発を目指そうとする取り組みが見られるようになってきた¹¹⁾。

以上のように、人工知能による人事データ分析は、人間よりも優れた公正中立的な判断を下すことができる、と期待されている。しかし、アルゴリズムに開発者等の先入観、誤解およびバイアス（偏り、偏見）が意図せずに入り込んだり、読み込んだデータの種類によって人間の予測しない誤った方向に学習が進むことが原因で、人工知能による分析は特定のグループに対して不利益ないし差別的な判断を行う恐れがある。それに加えて、分析手法によっては、自ら導出した判断や予測の結果に至る過程について、ブラックボックス化が進み、なぜそのような結果になったかについて、使用者も労働者も全くわからないといった事態も懸念される。これらの問題に対して、どのような法的対応を行うことができるのか、または、行うべきかが喫緊の課題となっている。

Ⅳ わが国の法的対応の現状

わが国では、使用者による労働者の個人情報の適正な取扱いを確保し、労働者の権利利益を保護するのは、「個人情報の保護に関する法律」（以下、個人情報

報保護法という)である¹²⁾。

同法では、個人情報の取扱いにあたって、まず、利用目的をできる限り特定しなければならず(第15条1項)、特定された利用目的の達成に必要な範囲を超えて、個人情報を取り扱ってはならないことになっている(第16条1項)。通常、使用者が労働者から各種の個人情報を取得し、それらを利用する目的は、人事労務管理、つまり、採用、配置、人事考課に始まり、作業管理、時間管理、賃金管理、安全衛生管理および職業訓練等のうちのいずれかである。使用者は、それらを労働者の個人情報の利用目的と特定して、その達成に必要な範囲内において、労働者の個人情報を取り扱うことができる。もちろん、使用者は、上述の人事労務管理上の諸利用目的のために、人工知能による人事データ分析で、労働者の個人情報を取り扱うことも許される。

また、同法では、個人情報の取得に際して、その利用目的を本人に通知し、または公表することが義務づけられている(第18条1項)。労働者から個人情報を取得するのに際して、使用者は、上述の人事労務管理上の諸利用目的を労働者本人に通知し、または公表しなければならない。しかし、使用者は、労働者に対しその「利用方法や手段」までも通知・公表することを義務づけられてはいない。したがって、人事労務管理のために人工知能による人事データ分析を行う場合も、使用者は、その実施の有無に始まり、どのようなアルゴリズムに基づき、取得・保管した諸データを分析し、いかなる結果や予測を導出するかについて、情報提供を行い、具体的に説明する義務はない。

このことから、使用者が、人工知能による人事データ分析を実施し、人事労務管理上の諸決定にその結果を利用したとしても、労働者に対しその実施等を自発的に知らせない限り、労働者は、その実施の有無、技術や手法、分析結果、その結果の利用法等を知る由もない。たとえ当該分析に基づく分析結果に、不正確さまたは誤りが含まれていても、労働者はその存在すら知らないのです。その訂正または削除を請求することはほとんど不可能である。さらに、不正確または誤った分析結果に基づく採否や処遇等の人事労務管理上の決定について、

その妥当性が疑われる場合でも、労働者が異議申立を行うこともきわめて困難である。

以上のように、現行の個人情報保護法の下では、使用者は、人事労務管理上の諸利用目的のために、人工知能による人事データ分析を実施し、人事労務管理上の諸決定にその分析結果を利用する限り、それらのことに関する何らの情報提供および説明を労働者に行わなくてもよい。他方、労働者は、人工知能による人事データ分析が実施されても、使用者からそのことを通知されなければ、その実施自体さえも全く知らず、不正確さまたは誤りを含みうるその分析結果に基づき、人事労務管理上の諸決定を甘受せざるをえない状況に置かれている。

このような状況を是正するための、法解釈による取り組みが見られる。たとえば、人工知能による人事データ分析を通じた精度の高い「予測」または「(情報の)付加」も、個人情報保護法上の「取得」に含まれるとし、その分析結果が要配慮個人情報に関わる場合には、要配慮個人情報自体を取得するときと同様に、当該分析を実施するのに際して、事前の本人同意を得る必要がある（第17条2項）、とする見解がある¹³⁾。しかし、「予測」も「付加」も、「取得」とは明らかに異なる行為であり、この法解釈には、明らかに無理がある。

このように、法解釈による一定の規制を加えようとする努力が見られるが、いずれにせよ、そのような努力のみでは、人工知能による人事データ分析をめぐる上述の諸問題を解決することは到底できない。やはり、人工知能による人事データ分析に対しては、立法などによるより積極的な法的対応が求められる、といえよう。

V 先進的な取り組みとしての EU 一般データ保護規則

1 序

人工知能によるデータ分析への法的対応の先進的な取り組みとして、EU 一般データ保護規則（General Data Protection Regulation - 以下、GDPR と

いう)がある。GDPRは、ビッグデータ分析、人工知能および機械学習によるプロファイリングや自動処理が個人の権利や自由に重大な影響を与える恐れがあるとして、個人の権利や自由、とりわけプライバシーを保護するため、プロファイリングなどの個人データの自動処理に対し、新たな規制を加えている¹⁴⁾。以下、その内容を紹介するとともに、批判的検討を加える。

2 個人データの自動処理のみに基づく決定の禁止と例外

GDPRでは、原則として、データ主体（識別されたまたは識別されうる自然人）は、自らについて法的効力を生じさせ、または、自らに同様の重大な影響を及ぼすプロファイリング¹⁵⁾などの自動処理のみに基づく決定¹⁶⁾に服さない権利を有する、とされる¹⁷⁾。「権利」という文言を用いているため、データ主体が積極的に求めた場合のみに同原則が適用されるという意味で受け止められてしまうが、同原則は、そのような意味ではなく、プロファイリングなどの自動処理のみに基づく決定に対する一般的な禁止を定めたものである。なお、本論の検討対象である、人工知能によるデータ分析を実施し、その他の要因を考慮することなく、その分析結果のみに基づき決定を下す場合にも、同原則が適用されうる。

ただし、①データ主体とデータ管理者との契約の締結または履行にとって必要な決定¹⁸⁾、②管理者が服し、また、データ主体の権利や自由および正当な利益を保護する適切な措置を命じる欧州連合または加盟国の法律によって正当化されている決定、③データ主体の明確な同意に基づく決定については、同原則は適用されない¹⁹⁾。このことから、労働者と使用者による労働契約の締結および履行のプロセスを前提にすると、使用者は、労働者の個人データに関するプロファイリングなどの自動処理を実施し、そのみに基づく人事労務管理上の決定を行うことを禁じられている、とはいえない²⁰⁾。ただし、許可されるのは、プライバシーへの侵害の少ない方法が利用できるかどうかを考慮して、目的達成のため、プライバシーへの侵害の少ない有効な方法が存在せず、自動処理に

基づく決定が必要である場合に限定される。たとえば、企業による求人募集に多数の応募者が殺到したため、有能な採用候補者を選抜するのに、個人データの自動処理による応募者の篩い分けが必要とされる場合などが、この例外にあたる²¹⁾。

3 データ保護影響評価

けれども、プロファイリングなどの自動処理、つまり、人工知能による人事データ分析を実施し、それに基づく人事労務管理上の決定を行おうとするときは、データ管理者である使用者は、まず事前に、予定の処理作業の個人データ保護への影響評価を実施することが求められている²²⁾。

データ保護影響評価（Data Protection Impact Assessment）とは、「データ処理を記述し、データ処理の必要性と比例性を評価し、個人データの処理に起因する自然人の権利と自由に対するリスクの管理を、（リスク評価とリスクに対処する措置の決定によって）支援するために設計されたプロセス」のことである²³⁾。データ保護影響評価の対象は、自然人の権利と自由にも高度のリスクをもたらす可能性があるデータ処理である²⁴⁾。このようなデータ処理について、データ管理者がその開始前に、データ保護責任者と処理者と一体となって、データ保護影響評価を確実に実施する責任を負っている。データ保護影響評価として、少なくとも、①想定される処理作業と処理目的を記述し、②処理の必要性および比例性を評価し、③データ主体の権利と自由に対する（処理の）リスクを評価したうえで、④リスクへの対処および本規則遵守の証明のための想定される措置を特定する。そして、これらのことを文書化し、データ処理を監視し定期的に見直すといったプロセスを繰り返すことが求められている²⁵⁾。

データ保護影響評価の実施に際しては、データ管理者は必要に応じて、データ主体またはその代表者の意見を求めなければならない²⁶⁾。他方、影響評価を公表する義務はないが、説明責任と透明性の観点から、データ管理者は少なくとも影響評価の概要や結論等の公表を検討する必要がある。影響評価の結果、

リスクを許容できる水準まで低減するのに十分な措置を見いだせない(つまり、残余リスクが高いままである)場合は、データ管理者は、リスクを低減するための措置の整備について監督機関と事前の協議を行わなければならない、とされる²⁷⁾。

4 適切な保護措置

次に、データ管理者である使用者は、データ主体である労働者の権利や自由および正当な利益を保護するのに適切な措置を講じなければならない。それには、少なくとも、管理者側の責任で人間を関与させることを求める権利、意見表明の権利、および、決定への異議申立の権利を保護する措置が含まれる²⁸⁾。ガイドラインにおいて、網羅的ではないが、適切な保護措置に関するグッドプラクティスの事例が示されている²⁹⁾。以下のとおりである。

- ・ 個人が差別されずに公正に取り扱われていることを確保するため、システムを点検すること。
- ・ 意図したように実際に動作していること、また、差別的、誤りのあるまたは不当な結果をもたらしていないことを証明するため、アルゴリズムを点検すること。
- ・ 独立の「第三者」による監査の際には、アルゴリズムの動作に関するすべての必要な情報を監査員に提供すること。
- ・ 第三者作成のアルゴリズムを利用するときは、その監査や点検がすでになされたことや、合意した基準にそれが対応していることについて、契約で保証を得ること。
- ・ プロファイルの作成または適用の際に使用されるプロファイルと個人データについて明確な保存期間を設定するなど、データ最小化のための措置をとること。
- ・ 匿名化または仮名化の技術を利用すること。
- ・ データ主体による意見表明や決定への異議申立を可能にする方法、および、

人間による関与のためのメカニズムを構築すること，など。

さらに，特別な種類の個人データ（いわゆるセンシティブ・データ）を伴う，プロファイリングなどの自動処理のみによる決定については，データ主体の明示的な同意がある場合，または，実質的な公共の利益を理由に必要とされる場合のうち，いずれかに該当しても，データ主体の権利や自由および正当な利益を保護する上述の適切な保護措置をデータ管理者が講じていなければ，許されない，とされる³⁰⁾。

5 情報提供義務とアクセス権

プロファイリングなどの自動処理，つまり人工知能による人事データ分析を行うのに際して，データ管理者である使用者は，公正かつ透明性のある処理を保証するため，その実施を始め，その意義や想定される結果，関連する論理（ロジック）について有意義な情報などを労働者に提供しなければならない³¹⁾。

ガイドラインによると，「意義や想定される結果」とは，自動処理に関する情報と，それに基づく決定がデータ主体に対しどのような影響を及ぼすかに関する情報のことである³²⁾。また，「関連する論理（ロジック）について有意義な情報」については，使用されるアルゴリズムの複雑な説明，または，アルゴリズムのすべての開示までも，必ずしも求められないが，決定に至るのに依拠する論理的根拠や基準など，データ主体が決定の理由を理解するのに十分に包括的な情報を提供しなければならない，とされる。たとえば，①プロファイリングまたは意思決定プロセスでこれまで用いたまたは今後用いるデータのカテゴリー，②これらのカテゴリーが関連すると考える理由，③分析で使用する統計学を含め，自動化された意思決定プロセスで用いるプロファイルの作成方法，④自動化された意思決定プロセスにこのプロファイルが関連する理由，⑤データ主体に関する決定にこのプロファイルが使用される方法などである³³⁾。

他方，データ主体である労働者は，プロファイリングなどの自動処理の実施の有無について，使用者から確認を得る権利を有するとともに，それが実施さ

れる場合、その意義や想定される結果、関連する論理（ロジック）について有意義な情報などにアクセスできる権利を有する³⁴⁾。また、労働者は、プロフィールなどの自動処理による自分についての結果が不正確または不完全であれば、遅滞なくそれを訂正し、または完全にしよう使用者に求めることができ、さらに、プロフィールなどの自動処理、または、それに基づく決定が適法に行われなかったときは、労働者は、それに関わる個人データの消去または取扱いの制限を使用者に求めることもできる³⁵⁾。

6 EU 一般データ保護規則への批判的検討

以上のことから、GDPR は、人工知能による人事データ分析の実施について、労働者の権利や自由および正当な利益を保護するため、データ保護影響評価と適切な保護措置を講ずることを使用者に義務づけるとともに、アクセス権の保障などデータ処理の透明性を高めることにより、労働者による自己情報へのコントロールを実現しようとしている。

しかし、最先端テクノロジーを利活用して、人工知能による人事データ分析がますます「ブラックボックス化」しつつある現状において、GDPR によるデータ処理の透明性の確保がきわめて困難になりつつあるのではなからうか³⁶⁾。GDPR は、データ処理およびそれに基づく決定についてデータ主体が理解できるように十分に包括的な情報を提供することをデータ管理者に求めているが、通常、データ主体である労働者は素人であり、いくら情報を受け取っても、それらを吟味してデータ処理の正確性またはそれに基づく結果の妥当性を確認できるであろうか。他方、GDPR は、使用されるアルゴリズムの複雑な説明、または、アルゴリズムのすべての開示までも求めているので、GDPR の求める範囲内の情報だけでは、専門家であっても、アルゴリズムにバイアス（偏り、偏見）が意図的に持ち込まれていないかどうか（または、意図せずに入り込んでいないかどうか）を解明することは困難である、といわれている³⁷⁾。さらに、人工知能による人事データ分析に用いられる方法によっては、なぜそのような

分析結果になったのかについて、そもそも、人間の言語でその論理（ロジック）を説明できない場合もありうる。それらに加えて、バイアスのかかったまたは差別的な分析結果であると疑われる場合でも、アクセス権の行使を通じて他者の個人データまでも取得できず、他者との比較を行うことができないため、その結果ないしそれに基づく決定に対し差別的であると異議申立を行うことはきわめて困難である³⁸⁾。このように、人工知能による人事データ分析については、データ主体である労働者は、GDPRの付与するアクセス権をもってしても、当該分析およびそれに基づく人事労務管理上の決定が客観的に正確かつ公正になされているかどうかについて、疑念を払拭することはできない。

とはいえ、GDPRは、人工知能による人事データ分析を行うのに際して、データ保護影響評価を事前に実施し、その後も定期的に繰り返して実施するよう、データ管理者である使用者に義務づけることにより、当該分析に基づく決定システム全体の絶えざる改善が図られ、当該分析に起因する労働者の権利や自由へのリスクが低減されることが期待されている。それと同時に、当該分析に起因するリスクから、労働者の権利や自由および正当な利益を保護するのに適切な措置の実施も使用者に義務づけられている。最も注目されるのは、ガイドラインにおいて、適切な保護措置のグッドプラクティスの事例として、データ管理者自らによるアルゴリズムやシステムの点検などにとどまらず、「独立の第三者による外部監査」までも示唆しているところである。

まだまだ人間にとって未知の領域である、人工知能によるデータ分析について、どのようなリスクが発生するのか、データ主体の権利や自由に、それがどのような影響を及ぼすのか、また、それに対しどのように対処すべきかなど、解らないことが多数存在する。それゆえに、データ主体一個人によって、問題の解決は図れない³⁹⁾。データ主体の権利や自由を保護するためには、組織的なリスク管理および対応が必要とされよう⁴⁰⁾。このような意味において、上述の「データ保護影響評価」や「適切な保護措置」は、データ管理者である使用者主導の取り組みであるため一定の限界があるが、人工知能によるデータ分析に

起因するリスクから労働者の権利や自由を保護するのに、大きな役割を果たすことが期待される⁴¹⁾。

VI わが国の法的規制についての基本的な方向性

以上のとおり、使用者が人工知能による人事データ分析を実施し、それに基づき人事労務管理上の決定を行うことは、労働者の権利や自由により重大な影響を及ぼす恐れがあるので、労働者の権利や自由を保護するため、当該データ分析の実施およびそれに基づく決定に対し法的規制を加える必要があることは明らかである。

今後検討されるべきわが国の法的規制の基本的な方向性について、上述のEUの先進的な取り組みを踏まえ示すと、以下のとおりである。

まず、使用者は、人工知能による人事データ分析の実施が労働契約の締結または履行にとって必要であり、なおかつ、労働者の権利や自由への侵害のより少ないその他の有効な方法が存在しないときに限り、労働者本人から自由意思に基づく明確な同意を得たうえで当該データ分析を実施し、その結果に基づき人事労務管理上の決定を行うのを許されるところと、

次に、使用者が当該データ分析を実施しようとする際は、労働者の権利や自由および正当な利益を保護するため、データ保護影響評価を行うとともに、適切な保護措置（人間の関与を求める権利、意見表明の権利および異議申立の権利の保障、アルゴリズムやシステムの点検、データの最小化など）を講ずるのを義務づけること、

さらに、当該データ分析を実施する際には、使用者は、その公正性と透明性を保証するため、その実施を始め、その意義や想定される結果、関連する論理（ロジック）について有意義な情報などを労働者に提供しなければならないとする一方、労働者には、当該データ分析の実施の有無を確認できる権利を始め、当該データ分析に関する上述の諸情報へのアクセス権、訂正権および処理制限権

などを付与し、また、労働組合など労働者代表には、（憲法第 28 条の団体交渉権のほか、）集団的アクセス権（使用者に情報提供を求める権利）や協議権などを付与すること⁴²⁾、

最後に、当該データ分析の開始後も定期的に、関係者からの意見を聴取しつつデータ保護影響評価を繰り返すとともに、適切な保護措置の一環として、第三者機関（データサイエンティストなど各種の専門家から構成された監査機関）による外部監査を実施するのを使用者に義務づけること、である。

人工知能による人事データ分析の実施に対するこのような法的規制は使用者にとって一定の負担となるであろう。しかし、労働者の権利や自由の保護の観点から、その分析結果の正確性や公正性およびそれに基づく人事労務管理上の決定の妥当性を確保するには、少なくとも、上述のような内容の法的規制が必要不可欠である、といえよう⁴³⁾。

Ⅶ 終わりに

新たなテクノロジーへの過剰な法的規制は、その発展を阻害するとの批判がある。しかし、そのテクノロジーが人間にとって未知の領域のものであり、なおかつ、人間に重大な影響を及ぼす恐れがある場合は、利害関係者の議論を通じて形成される一定の法的規制の枠内で、その開発および利活用が進められざるをえないであろう。本論の対象である人工知能によるデータ分析もまたしかりである。その有用性や利便性に目を奪われ、その潜在的リスクがなおざりにされないように、その開発および利活用が慎重に進められることを切に願う。長寿と繁栄を！

提出年月日：2019年9月5日

注

- 1) マイケル・ファーティック (著), デビッド・トンブソン (著), 中里京子 (訳)「勝手に選別される世界—ネットの『評判』がリアルを支配するとき, あなたの人生はどう変わるのか」ダイヤモンド社 (2015年), キャシー・オニール (著), 久保尚子 (訳)「あなたを支配し, 社会を破壊する, AI・ビッグデータの罠」インターシフト (2018年) など参照。
- 2) Jeffrey Dastin「焦点: アマゾンがAI採用打ち切り, 「女性差別」の欠陥露呈で」ロイター 2018年10月10日, 河鐘基「アマゾンの人材採用AIが『女性を差別した』理由を考えてみる」Forbes JAPAN 2018年10月16日, 金田捺「Amazonが人材採用システムを停止 —AIによる『差別』というダークサイドをどう乗り越えるか?」Ledge.ai 2018年10月23日 (<https://ledge.ai/ai-darkside/>) など参照。
- 3) 福原正大, 徳岡晃一郎『人工知能×ビッグデータが「人事」を変える』朝日新聞出版 (2016年), KPMG ジャパン編『ビッグデータ分析を経営に活かす』中央経済社 (2016年), 岩本隆『HRテクノロジー入門—AI・ビッグデータで加速する働き方改革と人事変革』ProFuture (2017年), 労務行政研究所『HRテクノロジーで人事が変わる—AI時代における人事のデータ分析・活用と法的リスク』労務行政 (2018年) 参照。
- 4) 山崎俊明『AI面接#採用』東京堂出版 (2017年), 労務行政・前掲注3) 書82頁以下参照。
- 5) 労務行政・前掲注3) 書126頁以下参照。
- 6) 労務行政・前掲注3) 書222頁以下参照。
- 7) Rebecca Greenfield, Riley Griffin, *Artificial Intelligence Is Coming for Hiring, and It Might Not Be That Bad - Even with all of its problems, AI is a step up from the notoriously biased recruiting process*, Bloomberg News, August 8, 2018. 塚越健司「アルゴリズムによる評価は人間よりも『公正』とは限らない 人工知能時代の『偏見』について」WEDGE Infinity (ウェッジ) 2018年3月23日。
- 8) Id. 塚越・前掲注7) 記事参照。
- 9) 河・前掲注2) 記事参照。
- 10) 渡部幹「『AIの判断は人種差別的』米調査に見るAI人事選考の危険性」ダイヤモンドオンライン 2017年11月22日 参照。Tom Simonite, *Artificial Intelligence Seeks An Ethical Conscience*, WIRED, 12.07.2017.
- 11) 「AIに潜む『性別・人種差別』の偏見 問題解決に何年もかかる理由」SankeiBiz 2018年3月6日 (ブルームバーグ by Dina Bass, Ellen Huet) 参照。
- 12) そのほかに, 使用者による応募者からの個人情報収集・保管・使用などについては, 職業安定法第5条の4 (求職者等の個人情報の取扱い) の規定およびその指針 (平成11年労働省告示第141号) がある。それらは, 個人情報の使用について, 収集目的に限られるとするのみで, 使用方法や手段についてまで規制を加えてはいない。
- 13) 山本龍彦編著『AIと憲法』日本経済新聞出版社 (2018年) 91-93頁, 労務行政・前掲注3) 書118頁以下参照。
- 14) A29 WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING AND PROFILING FOR THE PURPOSES OF REGULATION 2016/679 (Feb. 6, 2018) (hereinafter “A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-

MAKING”), at 5-6.

- 15) 本規則では、「プロファイリング」とは、自然人に関する一定の個人的側面を評価するための個人データの利用、特に、自然人の仕事上の業績、経済状況、健康、個人的嗜好、興味、信頼、言動、所在または活動に関する諸側面を分析または予測するための個人データの利用から構成される、あらゆる形態の個人データの自動処理、とされている（第4条4項）。
- 16) ガイドラインによると、「(自動処理) のみに基づく (決定)」とは、決定過程に人間による関与がないことを意味する。

個人データの自動処理がデータ主体に関する評価を下し、人間がそれとともにその他の要因を考慮し評価したうえで、最終的な決定を行う場合は、当該決定は、自動処理「のみに基づく」ものではない。他方、人間による関与を形式的に装ったとしても、自動処理で得られた評価を吟味したり、他の要因などを考慮したりすることもなく、常にその評価をそのまま決定とするならば、当該決定は、自動処理のみに基づくものとされる。人間による関与が認められるためには、決定の管理が、見せ掛けではなく、実質的なものであり、決定を変更できる権限と能力を有する者によって行われなければならない。A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 20-21.
- 17) 第22条1項。
- 18) 契約の締結または履行にとって、膨大な個人データを取り扱うのに、人間の手作業で処理することが非実際のまたは不可能であるため、自動処理に基づく決定が必要とされる場合がある。ガイドラインによると、「必要」かどうかについては、プライバシー侵害の少ない方法が利用できるかどうかを考慮して、目的達成のため、プライバシー侵害の少ない有効な方法が存在せず、自動処理に基づく決定が必要であるかどうかによって判断される。A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 23.
- 19) 第22条2項。
- 20) また、使用者は、交渉力・取引力の優位性を背景に、個人データの自動処理のみに基づく人事労務管理上の決定について「明確な同意」を労働者から取得して同原則の適用を回避しようとするかもしれない。しかし、第29条作業部会の意見（Opinion）は法的拘束力はないが、それによれば、労使間の使用従属関係に鑑みると、使用者による労働者の個人データの処理について、労働者が拒否しても不利益を受けない場合でない限り、労働者の同意はその法的根拠になりえない、とされる（A29WP, OPINION 2/2017 ON DATA PROCESSING AT WORK, WP249, Adopted on 8 June 2017）。
- 21) A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 23.
- 22) 第35条1項、3項。なお、データ保護影響評価の実施に関するGDPR上の義務を遵守しなかった場合、監督機関から罰金が科せられることがある（第83条）。
- 23) 第35条7項。また、データ保護影響評価は、管理者がGDPR上の義務を遵守することだけでなく、規則の遵守を確実にするのに適切な措置を講じているのを証明することにも役立つため、アカウントビリティーを果たすのに重要なツールである（第24条も参照）。つまり、データ保護影響評価は、コンプライアンスを確立し、証明するためのプロセスである。
- 24) 第35条1項。

- 25) A29WP, “GUIDELINES ON DATA PROTECTION IMPACT ASSESSMENT (DPIA) AND DETERMINING WHETHER PROCESSING IS “LIKELY TO RESULT IN A HIGH RISK” FOR THE PURPOSES OF REGULATION 2016/679” (revised on 4 October 2017), at 4.
- 26) 第35条9項。
- 27) 第36条。監督機関はデータ管理者に対し書面による助言を行うとともに、場合によっては、当該データ処理を一時的に制限し、または、恒久的に禁止することができる。
- 28) 第22条3項。
- 29) A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 32.
- 30) 第22条4項。A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 24.
- 31) 第13条2項、第14条2項。
- 32) A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 26.
- 33) A29WP, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 14, at 25-26, 31.
- 34) 第15条。
- 35) 第16条、第18条。
- 36) Lilian Edwards & Michael Veale, “*Slave to the Algorithm? Why a ‘Right to an Explanation’ Is Probably Not the Remedy You Are Looking For*”, 16 *Duke Law & Technology Review* 18, 67 (2017) (hereinafter “Edwards & Veale, *Slave to the Algorithm?*”); Lilian Edwards & Michael Veale, “*Enslaving the Algorithm: From a ‘Right to an Explanation’ to a ‘Right to Better Decisions’?*”, *IEEE Security & Privacy* (2018) 16(3), pp. 46-54 (hereinafter “Edwards & Veale, *Enslaving the Algorithm*”); Margot E. Kaminski, “*The Right to Explanation, Explained*”, U of Colorado Law Legal Studies Research Paper No. 18-24 (June 15, 2018), *Berkeley Technology Law Journal*, Vol. 34, No. 1, (2019).
- 37) Edwards & Veale, *Enslaving the Algorithm*, *supra* note 36, at 46-54.
- 38) Edwards & Veale, *Enslaving the Algorithm*, *supra* note 36, at 46-54.
- 39) データ主体が自分の代わりに異議申立等の権利を行使してもらうことを非営利団体等に委任するための「データ主体の代理」(第80条)や、データ主体の権利や自由を保護するための「監督機関」(第51条)については、本論では割愛する。
- 40) 「データ保護バイデザインおよび初期設定によるデータ保護」(第25条)や、「認証(制度)」(第42条)については、本論では割愛する。
- 41) Edwards & Veale, *Slave to the Algorithm?*, *supra* note 36, at 82; Edwards & Veale, *Enslaving the Algorithm*, *supra* note 36, at 46-54; Kaminski, *supra* note 36, at 23-24.
- 42) 人工知能によるバイアスのかかった評価やそれに基づく差別的決定から身を守るためには、各個人へのアクセス権(いわゆる自己情報コントロール権)の保障のみでは不十分である。というのも、そのような評価や決定から不利益ないし差別を受けるのは、一個人ではな

く特定の属性を有するグループであり、その不利益ないし差別を明らかにするには、各々のグループの構成員間の比較がまず必要であるからだ。また、このような比較を可能にするには、人工知能によるデータ分析およびそれに基づく決定について、使用者から十分な情報の提供と説明を受けることができる集団的アクセス権を労働組合等の労働者代表に付与する必要がある。それとともに、当該分析の正確性や公正性および当該決定の妥当性の確保に向けて、労働組合等の労働者代表が個々人に成り代わってより積極的にそれらに関与できるようにするために、協議権や異議申立権など（敢えて言うならば、集団的コントロール権）も付与される必要がある、といえよう。

- 43) もちろん、立法による法的規制のほか、労働契約上の附随義務としての人格権・プライバシー尊重義務または個人情報保護義務を根拠に、本文で列挙した各種の措置を使用者に義務づける、法解釈による取り組みも欠かせない。