

## 「法の支配」から「アルゴリズムの統治」へ

—— AIによる刑事司法の予測化・自動化における最低基本三原則：  
「公平性」「説明責任」「透明性」——

竹村 典良

### I プロローグ

今や、人間は「データ生産機械」になった。ネット検索、クレジットカード利用、ソーシャルメディアによる情報通信、医師の診断などにより、いつどこで何をしていたのか、誰あるいはどの機関とコミュニケーションしていたのか、何を好むのか、私たちにに関する情報の痕跡が残される。その上、広告主、データ・ブローカー、政府機関などは、私たちが日常生活で残したバン屑データを収集し、分析することができる。

情報通信技術は、個人や社会に利益をもたらすように使われることができるが、同時に、差別を助長し、不平等を悪化させるリスクも内包している。刑事司法の領域も例外ではなく、ビッグデータは、法執行に公平性をもたらす可能性を有するとともに、既存の偏見の再生産を助長する危険性も有している。

世界中の多数の国において、ビッグデータとアルゴリズムによる決定は法執行過程に埋め込まれている。これらの技術は司法をより迅速かつ平等にできると主張する。それらは、複雑で潜在的に偏見に満ちた裁量に基づく決定を、点数、数値、地図上のドットに変換する。

警察がビッグデータと予測分析に頼るようになると、警察取締りはより事前対策的で積極的になる。マイノリティーや移民がより高いリスク点数を取りやすくなり、再犯の危険性判断も高まる。その結果、これらの予測技術は、

マイノリティーや移民の犯罪者をより厳格に取り扱う、とするアルゴリズムの許可を裁判官に与えることになる。刑事司法の従事者が収集し、データベースに入力したデータに偏見が含まれていた場合、これらを使用した予測分析の結果も偏見に満ちたものとなる。予測司法技術は、既に刑事司法システムで作用している偏見と不完全を単純に再生産する高度の危険性を有している。そもそも、予測司法技術がどの程度実際に犯罪を減じ、予防し、予測するのは、現在分かっていない。

ビッグデータ、予測ソフトウェア、リスク評価アルゴリズムは、すでに刑事司法の風景・特徴を大きく変えているが、多数の問題を内包するという犠牲のもとに展開されている。ビッグデータとアルゴリズムによる決定は、正しい方向づけなくしては健全に発展しないがために、健全な懐疑的姿勢を保持し続けなければならない。

本稿は、第一に、A I、刑事司法、人権に関する問題状況を概観し、第二に、「機械司法」を通じてのアルゴリズム官僚主義による「安全統治」の実態を解明し、第三に、アルゴリズムによる統治を批判的に検討し、第四に、アルゴリズム統治の基本三原則（公平性、説明責任、透明性）について提案する。

## II A I、刑事司法、人権

A I による刑事司法の予測・自動化は、人権保障に関してどのような問題を引き起こしているのでしょうか。Aleš Završnik の研究「刑事司法、人工知能システム、人権」(Završnik, 2020) に基づいて、刑事司法の基本問題を批判的に検討する。

### I はじめに

A I システム、ビッグデータ分析、機械学習のような新しい技術は、犯罪のリスク評価や刑事司法システムの活動に重要な結果をもたらした。A I システムによる予測化・自動化は刑事司法の基本的問題についての再考を促す。判断理由の説明は何を意味するのか。どのようにしたら法的決定手続が透明

になるのか。誰が（半）自動決定の説明責任を負うべきか。A I の使用によって最終決定が促進された場合、どのように行為者集団の中で責任が分配されるべきであろうか。公正な裁判とは何か。刑事手続きの何らかの段階でA I システムが使われた場合、「法の適正手続き」は被告人に保障されないのであろうか（Završnik, 2020: 568）。

新しいA I システムが刑事司法の決定手続きにおいて使用されると、しばしば「ブラックボックス効果」が生まれる。結論に至る中間段階は、技術的な複雑性のために、人間の監視から隠されている。例えば、応用機械学習の複合領域は、どのように非監督学習あるいはアクティブ・ラーニングの新しい方法が人間の介入を回避するように行われているか、を示している（Završnik, 2020: 568; Pasquale, 2015; Leclerc *et al.*, 2020; Chan *et al.*, 2016; Brayne, 2018; Mehozay *et al.*, 2019）。

また、機械学習の作業がシステムを構築した研究者にさえも不透明である。この不透明性が機会学習を応用する多数の領域において問題にされないとしても、決定の説明可能性と論証の透明性が重要な価値を有する法的環境においてA I が使用される場合には透明性がなければならない。透明性と説明責任を欠く決定過程は、正当性があり非独裁性であるとは認められない。A I システムは本来的に不透明であるがために、刑事司法環境において使用される新しい道具は基本的な自由と矛盾する可能性がある（Završnik, 2020: 568; Završnik, 2018）。

## 2 予測化・自動化による犯罪統制の変化

### 1) 警察取締りの予測化・自動化

コンピュータを使用する「予測犯罪地図作成」(predictive crime mapping) 方法は、12年前に犯罪統制に導入された。予測的な「ビッグデータ」警察取締りの手段は一步前進した。A I の前進は、第一に、膨大な量のデータを理解し、散在するデータの集合から意味を推論することを可能にし、第二に、決定支援システムから主要な決定者への移行を表し、第三に、犯罪との闘いばかりでなく、社会全体の規制を目的とした（Završnik, 2020: 570; Saunders *et al.*, 2016）。

警察はA I を使用して、既遂犯罪を綿密に調査するばかりでなく、まだ起

きていない犯罪の予備段階の深部にも侵入した。予測化・自動化手段はいまだ起きていない犯罪の陰謀者を大量のデータから掘り出すことを想定し、「リスクのある」個人に焦点を当てる手段と「リスクのある」場所に焦点を当てる手段を区別した (Završnik, 2020: 570)。

しかしながら、管理監督された機械学習によるコンピュータの映像・パターン認識を利用する犯罪予防に関する研究は明らかに危険である。顔イメージからの自動的な犯罪性の推測に関する研究は、犯罪を推測するには単に 3 つの特徴 (口唇曲率、両眼内端間隔、鼻口交差角度) があるだけであるという結論を導いた。これらの研究者の潜在的仮説は、第一に、人間の容貌は内的特性に依存する、すなわち、人間には不変の核心がある。第二に、犯罪性は特定の人々の内的特性の表出であり、容貌を分析するだけで明らかにすることができる。第三に、これら 2 つの仮説が正しい場合には、刑事司法システムは犯罪性について信頼される決定をすることができる。実際、顔イメージから犯罪性を推測するソフトウェアは、犯罪とは何か、どのように犯罪が発見され、訴追され、判決が下されるのか、に関する根深い誤った考えを明らかにした。かつて嘲笑された 19 世紀の骨相学が「アルゴリズム骨相学」という衣をまとって 21 世紀に登場した。それは、犯罪に関する潜在的な偏見を正当化することができるとするが、容貌による有罪判断は重大なミスである (Završnik, 2020: 571; Eubanks, 2017)。

近い将来、犯罪統制に関し、身体に焦点を当てる方向にはさらなる進歩があることが示される。識別目的のための歩行パターン、態度、顔認証の分析から、感情の認識と心理状態の洞察のための表情と筆跡の分析まで (Završnik, 2020: 571; Gates, 2011)。

## 2) 刑事裁判の予測化・自動化

裁判所は、AI システムを使用し、被告人、保釈あるいはパロール (仮釈放) 中の犯罪者の再犯可能性と逃走可能性を算定する。最も分析され、議論されている事例は、そのようなソフトウェアが最も使われている米国におけるものである。米国の 21 の法管轄で使用されているアーノルド財団 (Arnold Foundation) アルゴリズムは、150 万件の刑事事例を使い、公判前段階における被告人の行動を予測する (Završnik, 2020: 571)。

136 万人の未決拘禁者を対象とする研究によれば、被疑者が逃走あるいは

再犯をするかどうかについて、コンピュータの方が人間の判事よりもより正確に予測することができた。しかしながら、このデータは説得力があるように見えるが、実際にはこの決定がより公正でないかもしれないことを検討することが重要である。個々の事件には、特異で、アルゴリズムで処理された40あるいはそれ以上のパラメーターを超え、熟考過程の結果に決定的であるかもしれない付加的な事実が常に存在する。したがって、永久に改善し続けることが必要である。裁判官はアルゴリズムが焦点を当てる変数よりもより広範な諸要素を視野に入れて判断しているであろう。AIシステムを使って何をしたいのか、何を「最適化」したいのかという問題がある。犯罪の減少は重要な目標であるが、刑事司法の唯一の目的ではなく、手続きの適正も同様に重要である（Završnik, 2020: 572）。

いくつかの欧州諸国では、自動決定システムを司法行政、とりわけ、判事への事件の割り振りに使用している。システムが基本的な透明性を欠くがゆえに、自動決定システムの独立したモニタリングと監査は不可能である。主要な関心は、システムが実際にランダムであるかどうか、そして、修理ができるかどうかである。より懸念されるのは、裁判管理目的で使用される自動決定システムが裁判官自身にも透明でないことである（Završnik, 2020: 572）。

### 3) 刑務所の予測化・自動化

新しい手段は有罪判決後の段階で様々な方法で使用されている。刑務所では、AIは安全の自動化や刑務所化からの社会復帰のために使用されることが多くなっている。中国で最も注目される犯罪者を収容する刑務所は、AIネットワークをインストールし、全受刑者を24時間体制で認識し、追跡し、異常が発生した場合に看守に警告する（Završnik, 2020: 572）。

これらのシステムは、処遇を通じて変化する犯罪者のニーズを確かめ、量刑手続きへの介入を監視するためにも使用される（Kehr *et al.*, 2017）。フィンランドの刑務所では、被収容者の訓練はAI訓練アルゴリズムを含んでいる。被収容者はユーザー研究において単純な問題を分類し回答する助けをする。この作業は刑務作業を組織する企業である Vainu に利益をもたらすと考えられており、受刑者も服役後の社会復帰に役立つ、仕事に関係する新しいスキルを身に付けることができる。米国における独居拘禁の危機に対処するために、アマゾンの Alexa のようなスマートアシスタントを受刑者の「拘

禁の仲間」として雇用するという、AI を用いる可能性について論じている研究者もいる。「仲間」の存在は一部の受刑者の心理的ストレスを軽減するかもしれないが、独居拘禁問題の「皮相部分」に焦点を当てることにより独居拘禁で悪化される危害に関する議論を隠蔽し、実際に独居拘禁刑罰政策の正当化に寄与する、と批判される (Završnik, 2020: 573)。

### 3 AI システムと法の邂逅

#### 1) 米国における法の適正手続きと AI システム

AI システムが刑事司法で最も使用されている米国において、ルーミス対ウィスコンシン事件判決におけるリスク評価アルゴリズム、「代替制裁のための矯正受刑者管理プロファイリング」(COMPAS: Correctional Offender Management Profiling for Alternative Sanctions) に関する決定は厳しかった。COMPAS アルゴリズムは、ルーミスを高い再犯リスクのために社会に対する高いリスクを示す者と識別し、第一審裁判所はパロール釈放請求を棄却した。上訴審において、ウィスコンシン州最高裁判所は、COMPAS アルゴリズムによる勧告はパロール釈放請求棄却の唯一の根拠でないがために、裁判所の決定はルーミスの適正手続権を侵害していない、と決定した。リスク評価アルゴリズムによる勧告の合憲性を確認する際に、ウィスコンシン州最高裁判所は「自動化の偏見」(automation bias) の影響の強さを否定した。裁判所は、下級審が提示されたアルゴリズム・リスク評価から逸脱した可能性があったと主張し、全てのアルゴリズム決定システムに含まれる偏見に関する社会心理及び人間・コンピュータ相互作用研究を無視した。それらの研究によれば、ひとたびハイテク手段による勧告が提示されると、人間の決定者がその勧告を拒否することは著しく耐えがたい負担となる。決定者は、勧告が不正確、不完全、誤りであることを知っていたとしても、自動化された勧告を中立的というよりもより積極的に評価する (Završnik, 2020: 573-574)。

カンサス州対ウォールズ事件判決 (2017 年) において、カンサス州上訴審裁判所は、ルーミス事件と反対の事実認定に至り、裁判所がプロベーション条件の決定に際して参考にした予測評価「行動レベル個人調査 [改訂版]」(Level of Service Inventory-Revised) 全体へのアクセスが被告人に許可されなければならない、と決定した。下級審裁判所は、被告人の LSI-R 評価

へのアクセスを拒否することにより、裁判所が使用した情報の正確性に対して被告人が異議申し立てをする機会を否定した。上訴審裁判所は、カンサス州対イースターリング事件判決に言及し、下級裁判所は LSI-R の全コピーを被告人に渡さなかったことにより、刑事手続きの量刑段階における適正手続きという憲法で保障された権利を奪った、と決定した（Završnik, 2020: 574）。

## 2) EUにおけるAIシステムの人権コンプライアンス

AIシステムは国家の義務に関係する人権に重大な影響を及ぼす。データ洪水がすべての社会領域に及び、アルゴリズム・システムが現代生活の多様な局面に浸透しつつあるがために、人権コンプライアンスはプライバシー・個人情報保護や差別のないこと・平等法の排他的領域だけの問題ではなくなっている。自動化システムは、銀行、保険、教育、労働など、多方面で導入されてきた。個人情報保護体制は、AIシステムの人権コンプライアンスに対する異議申し立てに対処するに十分ではない。インターネット仲介専門家委員会（The Committee of Experts on Internet Intermediaries [MSI-NET]）が欧州評議会で認めたように、人権の含意は多様である。自動化手続技術やアルゴリズムの使用によって影響を受ける人権は、1）公平な裁判と適正手続きの権利、2）プライバシー・個人情報保護、3）表現の自由、4）効果的な救済を受ける権利、5）差別の禁止、6）社会権と公共サービスへのアクセス、などである。その上、基本的自由は相互依存・相互関係にあるがために、アルゴリズム技術が教育、社会福祉、民主主義、法システムにおいて使用されると、全ての人権が潜在的な影響を受ける。社会システムや領域で使用されるAIの開発は、国家介入に抗する保護シールドとしての人権概念そのものを破壊することがある（Završnik, 2020: 574-575; The Committee of Experts on Internet Intermediaries [MSI-NET], 2017; Brkan, 2017; 渡邊, 2018）。

## 4 小結

アナログ世界における法の支配原則を尊重する人権コンプライアンス・システムをどのように設計するか。アルゴリズムの蔓延は政策立案者の関心を呼び起こした。アルゴリズムが基本的人権に及ぼす影響とアルゴリズムの責任をどのように構成するか、に関心が寄せられている。ヨーロッパでは、

2018 年末に、欧州評議会の司法の効率のための欧州委員会 (Council Europe's European Commission for the Efficiency of Justice [CEPEJ]) は、アルゴリズムが人権に及ぼすリスクを減じるために、「司法システムにおける AI の使用に関する欧州憲章」(The European Charter on the Use of AI in Judicial Systems) を採択し (The Committee of Experts on Internet Intermediaries [MSI-NET], 2017; European Commission for the Efficiency of Justice [CEPEJ], 2018)。同様の関心は世界の他の地域でも見られる。米国では、ニューヨーク市評議会が、アルゴリズム決定の透明性に関する法律を世界で初めて可決した。この法は、自治体機関によって使用されるアルゴリズムの公平性と妥当性をモニターするタスクフォースを設置した (Završnik, 2020: 579)。

刑事司法と警察取締りにおける AI の使用は、いくつかの刑事手続き上の権利 (無罪推定、公平な裁判を受ける権利 (武器の対等、証人に対する反対尋問権)、独立かつ公平な裁判を受ける権利 (裁判官の任意選択ほか)、無差別・平等原則、正当性原則) に潜在的な影響を及ぼし、既存の証拠規則を曖昧にする。法律家を含む社会学者は、コンピュータ・データ科学者とともに、人権コンプライアンスを作成し、事例研究を分析することにより、将来において使用される新しいシステムの人権コンプライアンスに多大な利益をもたらすであろう (Završnik, 2020: 579-580)。

### Ⅲ 「機械司法」：アルゴリズム官僚主義による「安全の統治」

アルゴリズムによって駆動されるリスク評価ツールが予測警察取締りや予測司法に使用されると、刑事司法が機械化され、アルゴリズム官僚主義による「安全の統治」が展開される。Rik Peeters と Marc Schuilenburg による研究「機械司法：アルゴリズムの官僚主義による安全の統治」(Peeters *et al.*, 2018) に基づいて、予測警察取締り、予測司法、機械司法の問題について、批判的に検討する。

#### 1 アルゴリズム予想：監視資本主義と安全の統治

近年における情報通信技術の急激な発展はアルゴリズムを使う新しい分野を切り開いた。新しい技術的実行性により、行動パターンを発見することを目的として、膨大な量のデータを迅速に分析することができる。アルゴリズムによる行動予測は、Zuboffが「監視資本主義」と名付けた強力なビジネスモデルである（Peeters *et al.*, 2018: 270; Zuboff, 2015; Zuboff, 2019）。

現代情報通信技術の監視能力は刑事司法でも利用されて来た。「市民の得点付けとランク付けの激増」（Harcourt, 2015: 205）は、人の犯罪傾向や被告人の再犯傾向の予測まで拡張されている。ここでの目的は、財政的というよりは統制志向である。刑事司法における情報技術の利用は新しいものではなく、過去数十年間にアルゴリズム駆動リスクシステムの数は著しく増加した。Harcourt（2007）によれば、プロファイルと予測は安全と予防を希求する際の第二の天性となった。これらリスクシステムの中核的な要素は行動を優先することである。予防と処罰の目的のために、個人あるいは特定集団の行動がモニターされ、予測され、必要な場合には、介入に従わせる。新しい形態のデータ分析が監視の機会を拡張するばかりでなく、新たな段階が加わるということを理解することが重要である。アルゴリズムを利用し、リスク評価装置によって知覚されたリスクに従って、犯罪が発生する高度の可能性がある場所、個人、集団を選出することが定量化される。個人の状態を決定するのではなく、個人あるいは集団が示すリスクを決定するために、データが利用される。警察は、犯罪歴、性別、年齢、低レベルの学歴、非定住、失業のような犯罪行動のリスク要因を多数備える個人をターゲットにすることができる（Peeters *et al.*, 2018: 270-271）。

刑事司法におけるアルゴリズムの利用は、一部は技術発展機会の増加の結果であり、また、一部は、政府がリスクを予測し回避するために尽力し、犯罪統制を超えてより広範な安全コンテキストにまで関心を拡大した結果である（Schuilenburg, 2015）。これらの発展が集中し、未踏の領域に情報技術とアルゴリズム予測を応用する強力な動機が生み出された（Peeters *et al.*, 2018: 271）。

## 2 予測機械刑事司法システム

### 1) 予測警察取締り

「予測アルゴリズム」と「犯罪探知」の利用は、現実の表象を単なるデータのフローから未来への投影に変えた。犯罪統計からソーシャルメディアのテキストとイメージの採掘まで、多様な情報源から得られた「ビッグデータ」を寄せ集め、自動分析することにより、犯罪パターンと被疑者、地理的位置、ネットワーク化された性質を明確にする診断結果が導かれる。これらのアルゴリズム駆動リスク評価システムの究極目的は、予防活動によって予測された結果を変えることである (Perry *et al.*, 2013)。概念的には、予測警察取締りは、知識主導警察取締り (intelligence-led policing)、データ駆動警察取締り (data-driven policing)、リスクに基礎を置いた警察取締り (risk-based policing) のような他のデータ注入警察取締り (data-infused policing) と密接な関係にある。他と異なるのは、単なる既存のデータからの推測ではなく、犯罪パターンの変化の予測もできるモデルの利用が含まれ、未来への明確な投影ができることである (Bennett Moses *et al.*, 2016; Peeters *et al.*, 2018: 271)。

予測警察取締りは、4つのレベルにおいて、将来の犯罪を予測するために使用することができる。1) 犯罪が起りやすい場所 (しばしば、「ホットスポット」と呼ばれる)、2) 犯罪を起こしやすい人物、3) 犯罪が起りやすい時間、4) 場所、人物、時間の組み合わせ、である。基本的な仮定は、「特徴のあるプロファイルを持つ犯罪者が同類型の犯罪を同じ場所と時間に犯す傾向があるがために、窃盗や強盗のような犯罪の大部分は予測することができる」ということである。統計学的手法により、逮捕の件数・種別、犯行場所に関する既存の警察記録が使用され、犯罪行為のパターンが認識され、予防的アプローチ、集中的で特定された迅速な対応を促進することができる (Peeters *et al.*, 2018: 272; Schuilenburg, 2015)。

予測警察取締りの興味深い情報源は、ソーシャルメディアのユーザーによって自発的に露出された社会的ビッグデータである。政府はこれらのデータを使用して、将来の犯罪行動の兆しとなるおそれのある、過激集団や無秩序に関わる人物に対する忠誠の表明を厳しく取り締まる。類似の戦略は、ツイッターの掲示を分析し、犯罪パターンを地図化することである。例えば、公共の場所における小さな無作法な行為に関するツイートをビッグデータとして分析することにより、特定の地域社会における「割れ窓理論」を確認する

ことに役立つ（Peeters *et al.*, 2018: 272; Williams *et al.*, 2017; Harcourt, 2015）。

## 2) 予測司法

アルゴリズムは、警察実務ばかりでなく、刑事司法システムの管理運営も変形する。単なる応報に代わり次第に功利主義的になっている行刑学の背景に反して、データは裁判所によって使用され、個人の「将来的危険性」が予測される（Berk *et al.*, 2013）。アルゴリズムが刑事司法システムに適用され、被告人の逃走リスクや公共の安全に対する脅威が評価される。統計数理予測方法、再犯リスク・アルゴリズム、リスク評価ツールは、ある者が高・中・低いどれかのレベルのリスクを示すかを推測するために、被告人の犯罪歴、社会人口統計的特性に関する変数を使用する装置である。主たる変数は過去の犯罪事件記録に基づいて認定される。その目的は、証拠に基づくリスク評価により、量刑、保釈決定、社会復帰プログラムを改善することである。より具体的には、アルゴリズムは、公平性（全リスクの平等な扱い）、透明性（客観化されたスコア）、効率性（準備期間）を高めることを目的とする（Peeters *et al.*, 2018: 272-273; Goel *et al.*, 2016）。

現在、刑事司法と犯罪精神医学において 200 以上のリスク評価ツールを手に入れることができ、量刑、パロール審査、釈放後モニタリングなど広範囲に利用されている（Douglas *et al.*, 2017: 134）。とりわけ、米国では、特別な変数セットとともに、3つの主要なシステム（COMPAS、PSA、LSI-R）が用いられている。興味深いのは、民間企業で働く統計専門家、数学者、コンピュータ科学者、ソフトウェア・エンジニアがこれらのシステムを開発し、アルゴリズムの作業結果には民間の著作権があり、公表されないことである（Peeters *et al.*, 2018: 273）

## 3) 機械司法

多数の研究者がアルゴリズム駆動リスク評価の危険性と可能性を指摘してきたが、これらのツールが法秩序維持の任にある公務員の行動を統制するゲームのルールにどのような影響を及ぼすかについてほとんど知られていない。情報通信技術の使用は街路レベルにおける決定者の裁量の余地を減少させる傾向にある。街路レベルの公務員はデータとアルゴリズムに依存するようになる。アルゴリズム駆動リスク評価は、警察取締りと刑事司法システムを専

門家の評価から機械のような手続き分類に変化させる。アルゴリズム駆動リスク・ツールの機械学習能力は、決定過程は絶えず流動的であるとするが、これらのツールの使用はより硬直し標準化した行動に向かう (Peeters *et al.*, 2018: 274-275)。

アルゴリズムが導入される前からリスク評価は行われていた。例えば、保釈、パロール、プロベーションの決定は、再犯や裁判所への出頭と関係してきた。しかしながら、これらのリスク予測は組織化・規則化されておらず、裁判官や精神医学者の専門的な判断に任されていた。これとは対照的に、情報通信技術が官僚的領域において標準化され、社会的複雑性があらかじめ定められた点数に減じられ、アルゴリズム駆動リスク評価は機械論理に従って機能するようになる。一度アルゴリズムがデザインされると、人間の介入がなく、結果が自動的に生み出され、個別化司法が危険に晒される。アルゴリズムがデータ処理のための「デジタルレシピ」として作用するような決定方法の変化による専門的な裁量の減少は下記の事実によってますます促進される。現在利用可能なリスク評価ツールは民間企業によって作成され、著作権が保持されるがため、徹底評価のために閉ざされたブラックボックス化する (Bennett Moses *et al.*, 2016)。これにより、警察の配置、量刑、プロベーション、保釈に関する決定のための説明責任は著しく疑わしくなる (Peeters *et al.*, 2018: 275; Brucato, 2015; Hamilton, 2015)。

これに対して、予測的警察取締りや予測司法のシステムに従事する者は、アルゴリズムを無視し、個別事例に即した自身の評価をする自由がある、という批判が考えられる。確かに、この批判は形式的には正しいが、官僚主義的实践においては、専門的裁量を自己抑制する傾向が見られる。アルゴリズムは、社会的複雑性をあらかじめ定められたカテゴリーやリスク要因に縮小することの探求と、効率性と客観性の探求を結合する。その結果、アルゴリズムの結果は、個別事例の個々の評価よりもよりインプットとデザインによって決定される (Peeters *et al.*, 2018: 275)。

### 3 擁護者と批判者

警察実務と刑事司法システムにおける決定を自動化するアルゴリズム駆動リスク評価の利用を擁護する者は、統計数理的予測ツールの使用を「科学化、

節約化、民主化」と考える（Smith *et al.*, 2017: 261）。第一に、科学技術の進歩と発展はリスク・危険性評価への新しいアプローチに扉を開く。これらの新しいツールは盲点を発見し複雑な社会問題を解決することができる、という強い信念がある。第二に、統計数理予測ツールは政府の効率性を高める。法執行機関は、犯罪遂行リスクの高い市民を取り調べることにより、これまでと同じ資源でより多くの犯罪を発見することができる。刑罰執行機関は、将来再犯を犯す可能性が高い市民を隔離することにより、犯罪を減らすことができる（Harcourt, 2007: 3）。第三に、これらの新しいツールは価値中立な方法で刑事司法に関する知識を生み出し鑑定することができる、と一般に考えられている。それらは、時間と費用を節約するばかりでなく、人間の偏見を回避し、差別と人種差別的な量刑実務を減じることにより、量刑をより説明可能なものとする（Christin *et al.*, 2015; Peeters *et al.*, 2018: 274）。

これに対して、アルゴリズム駆動リスク評価に対して批判的な視点は、アルゴリズムによる犯罪・再犯予測の信頼性に関する深刻な関心を示す。「全世界の完全な監視」は完全無欠のデータ収集が必要であるがため、完全な予測は不可能である。その上、信頼性のある予測をするのに必要とされる膨大な量のデータは、批判的な信号を可能性のある相関関係の霧の中に隠すことができる。モデルは常に関係するデータの選択と標本化を必要とするがために、必然的に歪曲と歴史パターンの特殊な投影に終わる機会が増加する。既存の警察統計の先入観（通報されない犯罪、前科など）はアルゴリズムのデザインに反映し、すでに過剰な警察取締りが行われている極貧・非白人居住地域が強調される。非白人被告人はより多くの機会に「誤った肯定」（高いリスクと評価されるが再犯しない）を、白人被告人は「誤った否定」（低いリスクと評価されるが再犯する）を受けることが明らかである（Peeters *et al.*, 2018: 274; Rehavi *et al.*, 2014）。

統計数理予測ツールは社会復帰アプローチを処罰に変える、との指摘がある。被告人のニーズよりも社会に対するリスクに重点が置かれ、リスク評価ツールは低リスクの被告人を排除するのではなく処罰を惹起する。これはリスク評価ツールの意図せぬ結果であるかもしれないが、これらのツールは政治的・イデオロギーの影響を受けている、と批判される。過去数十年の犯罪政策における「厳格」イデオロギーと懲罰転回言説に代わり、統計数理転回

は、客観性と経験主義を装い、処罰アプローチにとってよりクールで微妙な陰影を持つ基礎を提示する (Peeters *et al.*, 2018: 274)。

#### 4 小結

最後に、刑事司法システムにおいて、アルゴリズムが使用され、システムが予測機械化する危険性を減じる方法を議論しなければならない。アルゴリズムの説明責任を確実にするために、政府による規制あるいは独立した監督機関を創設する必要性、また、アルゴリズムのデザイン、入力データ、潜在的偏見を評価する必要性、が提案される。また、アルゴリズムのデザイン段階において、多様性 (多様な背景を持つ人々) を高めることも提案される。アルゴリズム中心の決定を回避するために、個々の事例における人間による評価を強調するメカニズムが必要である。市民の法的権限を強化し、アルゴリズムにおける変数や決定者が用いる個人データに関する市民の地位を高め、不正なアルゴリズム決定を訂正する権利を導入することも考えられる。刑事司法システムの予測機械化に直面し、人間による決定と機械による決定のバランスを回復するための手段が市民に与えられなければならない。

### IV アルゴリズム統治、監視資本主義、情報市民化

アルゴリズムの統治は、デジタル技術が特別な方法で社会秩序を生成するデジタル社会の登場に関する議論における重要概念である。Rob Kitchin による研究「アルゴリズムに関する批判的思考と研究」(Kitchin, 2017) に基づいて、アルゴリズム統治が複合的で偶発的で競争的であり、多様なコンテキスト・領域で多様な形態を取り、利益、権力、抵抗によって形成される状況を批判的に検討する。

また、監視資本主義は、情報通信技術による社会支配において必然的に出現する資本主義の一形態であり、「ビッグ・アザー (大いなる他者)」(Big Other) による監視という特殊な社会統制が出現している。Shoshana Zuboff による研究「ビッグ・アザー: 監視資本主義と情報市民化の可能性」(Zuboff, 2015) に基づいて、「監視資本主義」(surveillance capitalism) の問題状況

を分析し、そこからの脱出の可能性を探求する。

## 1 はじめに

日常生活のより多くの局面がソフトウェアによって動作可能となる技術によって、媒介され、増加され、生産され、規制される。ソフトウェアは、基本的に、アルゴリズム（指示・データを処理して結果を生み出すように構成された一連の段階）によって構成される。アルゴリズムに関して、4つの議論が展開される。第一に、社会・経済生活の形成において増大する重要性を鑑み、アルゴリズムとその業績について批判的かつ経験的に注意を向ける必要性が迫っている。第二に、アルゴリズムは、技術的、コンピュータ計算的、数学的、政治的、文化的、経済的、コンテキスト的、物体的、哲学的、倫理的など、多数の方法で捉えることができるが、偶然的、存在生成的、遂行的な性質で、より広範な社会技術的集合体に埋め込まれた存在として、最もよく理解される。第三に、アルゴリズムに関する研究を妨げる3つの難題（組織的な記述へのアクセスの獲得、異種混合で広範なシステムに組み込まれている、作動がコンテキストの中で偶然的に展開する）があり、実践的・認識的注意が必要である。第四に、アルゴリズムの構成と作動は多様な方法で経験的に研究でき、それぞれに系統的に評価する必要のある長所と短所がある（Kitchin, 2017: 14; Veale, 2019; Cino, 2018; Hanna-Moffat, 2019）。

現在、私たちは、アルゴリズムが権力行使において常に増大する役割を果たす広範な「アルゴリズム統治」の時代に突入しつつあり、アルゴリズム統治により、社会の規律・統制が自動化され、資本蓄積の効果が増大する。しかしながら、アルゴリズムの権力行使の方法が不明確である。アルゴリズムはデータに基づいて重要な決定を行うように組み込まれているが、著しく「ブラックボックス化」しており、疑義を挟む余地がないがために、明確性が欠如する。（Kitchin, 2017: 15; Pasquale, 2015）。

## 2 アルゴリズムの批判的考察

### 1) アルゴリズムの作成

批判的研究者によれば、アルゴリズムは、客観性、公平性、信頼性、正当性、のいずれの性質もない。コードは、理論的・数学的でなく、きわめて社

会的・政治的局面があり、あらゆる種類の決定、政治、イデオロギー、ハードウェアと基礎構造の実体に枠づけられ形成されている。プログラマーは、地方慣習、文化、知識、コンテキストから距離を置き、分離し、公平に作動させ、独立することにより、高度の機械的客観性を保持しようと努めるが、タスク・プロセス・計算をアルゴリズムに変換する際に、これらから完全に逃れることができない。彼らは、入手可能な資料および訓練データの選択と質、スタンダード、プロトコル、法の要請、ハードウェア、プラットフォーム、バンド幅、言語に関係する選択と条件、などの要因から逃れることもできない。したがって、現実には、アルゴリズムの作成において、高度の専門的知識、判断、選択、制約など、多数の要因が作用する。その上、アルゴリズムは、しばしば中立とは全くかけ離れた目的のために作られる（価値と資本の創出、一定の方法で行動と構造の志向を推進する、人々を確認し、区別し、分類する）(Kitchin, 2017: 17-18)。

以上のように、アルゴリズムの作成は、文化的、歴史的、制度的に位置づけられた社会的・実体的実践の交差点で行われる。アルゴリズムは開発され採用される条件から切り離すことはできない。アルゴリズムは、関係的、偶然的、コンテキスト的な性質を有し、社会技術的な集合体という広範なコンテキストの枠組みに組み込まれたものとして理解されなければならない。このような視点からすれば、アルゴリズムは広範な装置の中の一要素であり、技術的・客観的・公平な形態の知識あるいはオペレーション様式として理解されてはならない (Kitchin, 2017: 18)。

## 2) アルゴリズムのパフォーマンス

アルゴリズムの作用、影響、権力について検討することが必要である。アルゴリズムが中立・公平な知識の表現でないように、その作用は無感覚で非政治的ではない。アルゴリズムは、人々、プロセス、場所について、検索、整理、分類、カテゴリー分け、グループ分け、照合、分析、プロファイル、モデル化、シミュレーション、可視化、規制する。アルゴリズムは、私たちの世界理解の方法を形成し、ソフトウェアを作動させて世界を作り、深淵な結果をもたらす。この意味において、アルゴリズムは事態を惹起するがために遂行的である。これらのアルゴリズムの作成者は、偏向的あるいは自己奉仕的な媒介者を置き換えあるいはその役割を減じ、決定から主体を取り除く、

と論じるが、コンピュータ計算は、分類と差異的処遇のプロセスを深化・加速し、伝統的な病理学を改善するよりも具体化した（Kitchin, 2017: 18-19）。

アルゴリズムは、全く価値中立的ではなく、権力と知の体制を構築し適用し、その使用は標準化を暗示する。アルゴリズムは、強制、規律、規制、統制し、人間・動物・客体が相互作用し、多様なシステムを経験する方法をガイドし再構成するために用いられる。アルゴリズムは、アルゴリズム権力（algorithmic authority）、アルゴリズム統治（algorithmic governance）であることを主張し、表現する。その結果、社会は生活の指針となるアルゴリズム的生成的規則を持つ。そのような規則は、コンピュータ操作に埋め込まれ、アルゴリズムによる権力の表明であり、多様な現実を生み出す「仮想」である。これらの規則は、圧縮され、隠され、それらと視覚的に遭遇することはできない。それらは資本権力が作用する通路である（Kitchin, 2017: 19; Lash, 2007）。

### 3) アルゴリズムの存在生成性、遂行性、偶発性

アルゴリズムの影響あるいは権力は、以下の3つの理由から、常に線形的あるいは常に予測的であるのではない。第一に、アルゴリズムはその作用を媒介し屈折させる関係の広範なネットワークの一部として作動する。第二に、アルゴリズムのパフォーマンスは副次効果と意図せぬ結果をもたらすことがあり、世話・監督を受けずに、予期せぬ行動をすることがある。第三に、アルゴリズムはバグ（欠陥）あるいは誤ったコード化のために偏りと誤りをすることがある。一度コンピュータ操作が公開されると、ユーザーによってあらゆる種類の代替的方法で技術が埋め込まれ、多様な手段のために使用され、アルゴリズムの意図が批判され、覆され、再構成される。この意味において、アルゴリズムはプログラマーが作成したもの、あるいは、特定のインプットに基づいて作成された影響ではなく、日々ユーザーが作成し続ける（Kitchin, 2017: 19）。

アルゴリズムは、時に結果を予測することが容易でなく、予期せぬ効果を生み出すという意味において、しばしば「統制不可能」である。アルゴリズムの作動と影響を理解するためには、状況、時空を横断するコンテクスト的偶発的展開に敏感でなければならない。これが実務において意味するのは、アルゴリズムとの単純で限定的な関わりはあらゆる場面で推定できず、ひと

まとまりの比較事例研究、あるいは、多様な条件のもとで作用する同一のアルゴリズムについての一連の実験が必要である (Kitchin, 2017: 21-22)。

### 3 監視資本主義と情報市民化

Zuboff は、「ビッグ・アザー (大いなる他者)」というキーワードで表現される「監視資本主義」の問題状況を分析し、そこからの脱出の可能性を探求する。ネットワーク化された領域において蓄積される新しいロジック「監視資本主義」が台頭し、これに伴い「情報市民化」(information civilization) が胎動し始めている。コンピュータ・アーキテクチャー (コンピュータ・システムの構成・設計思想) は、拡散し、異議申立の余地のない新しい権力「ビッグ・アザー (大いなる他者)」を生み出す。それは、採取、商品化、統制の予期・解読不可能なメカニズムで構成されている。このメカニズムは、人々を自身の行動から追放し、行動予測・変容の新しい市場を生み出す (Zuboff, 2015: 75)。

現実生活をめぐる、ストレスに充ち、競争的で、階層化された闘争において必要な資源として、人々は急速に新しい情報コミュニケーション・ツールに依存するようになった。こうして、新しいツール、ネットワーク、アプリケーション、プラットフォーム、メディアは、社会参加のために必要不可欠になった。データ仲買業務、データ分析、データ採掘、データ専門家、理解不能なキャッシュ・フロー、強力なネットワーク効果、国家共同、ハイパー規模の財産、未曾有の情報権力の集中のような制度化された事実の増強は、不可避という圧倒的な感覚を生み出した。これらの発展は、完全に制度化された新しい蓄積ロジック「監視資本主義」のための基盤となった。この新しい体制において、グローバルなコンピュータ・メディア・アーキテクチャーは、組織に結び付けられた電子テキストを世界中に伸びる知的有機組織「ビッグ・アザー」に変えた。新しい征服 (服従) の可能性が生み出され、その革新的なロジックは予期不可能で解読不可能な採掘・統制メカニズムの上で成長している (Zuboff, 2015: 85)。

自動化され、遍在するビッグ・アザーのアーキテクチャー、監視資産におけるその展開、浸透的監視としてのその機能は、この蓄積ロジックの他の驚くべき新しい特徴に新しい光を当てる。それは、市場と民主主義の歴史的な

関係を傷つけ、人々に無関心で、著しく離れた組織を形成する。監視資本主義は、人々と資本家が雇用と消費のために相互を必要とした伝統的な相互依存関係の影響を受けない。この新しいモデルでは、人々はデータ採取のターゲットとなる。この社会との隔絶は監視資本主義の反民主主義的な特徴のいま一つの局面である。監視資本主義の下では、民主主義は監視の収益を脅かすために、もはや繁栄の手段としては機能しない（Zuboff, 2015: 86）。

最後に、監視資本主義は、現代におけるヘゲモニック（指導的）な蓄積ロジックなのか、それとも新しく生まれつつある他の情報を基盤とする市場形態にヘゲモニーを移譲する革命的袋小路となるのであろうか、将来の代替的な軌道はこれらの形態の争いの結果に左右される。情報民主化が実現するか否かはこれらの問題への解答にかかっている（Zuboff, 2015: 86; Zuboff, 2019）

#### 4 小結

世界中の人々は、日常的に、通信、公共事業、輸送、デジタル機器の電源において作用するソフトウェアに埋め込まれた何百というアルゴリズムと接触している。これらのアルゴリズムは、破壊と変形を引き起こし、システムの作用を再構成し、新しい形態のアルゴリズム統治を実行に移し、新たな形態の資本蓄積を可能にする。その拡散、活用、自動化への投資力は著しいが、技術的な視点からアルゴリズムにアプローチする膨大な研究と比べ、アルゴリズムを批判的に検討するものは殆ど無い。コンピュータ計算が豊富な世界において、アルゴリズムがどのように思考し、知的に関わるのか、についてのこのような不均衡はいささか驚きである。社会科学と人文科学において、アルゴリズムとアルゴリズム統制の形態について批判的な検討をすることが喫緊の課題である（Kitchin, 2017: 26）。

アルゴリズム統治は、一つのことではなく、多様な領域とコンテキストにおいて多様な形態を取り、利益、権力、抵抗によって形成される。アルゴリズム統治は複合的で偶発的で競争的である。アルゴリズム統治の概念は、特定の方法で社会を秩序付け規制する広範な社会技術的实践を含む。この概念の利点の一つは、多様な現象、言説、研究領域を結合し、新生のデジタル社会の主要な論争と異議申し立てを明確にすることである。偏見と公平、透明

性、人間機関は取り組まなければならない重要な問題である。アルゴリズム統治は、秩序立て、規制、行動変容、管理支配、最適化、参加など、多くの顔を持つ。研究領域に基づいて、不可解、価値・利益の登録、能力と効率性、権力不均衡、社会的包摂、排除、競争、応答性、参加、共同創造と負担によって特徴づけられる。大部分の者にとって、アルゴリズム化とデータ化により、統治は権力的で、侵入的で、浸透的であるが、これと異なり、統治は包摂的で応答的で社会的多様性を許容すると強調する者もある。アルゴリズム統治は複合的で、社会的経済的的政治的コンテキストに関して偶発的である (Katzenbach *et al.*, 2019: 11)。

## V 刑事司法におけるアルゴリズム使用最低基本三原則：「公平性」 「説明責任」「透明性」

A I、予測評価アルゴリズムを刑事司法システムで使用するために必要不可欠な最低基本三原則として、「公平性」「説明責任」「透明性」を確立し、実践することが重要である。Danielle Kehl, Priscilla Guo, Samuel Kessler による研究「刑事司法システムにおけるアルゴリズム：量刑におけるリスク評価の使用についての評価」(Kehl *et al.*, 2017) に基づいて、A I 使用基本原則について検討する。

### 1 はじめに

過去数年にわたり、法学者、政策立案者、実務者ほかは、A I、機械学習、ビッグデータ、予測のためのソフトウェアを刑事司法の領域で使用することの倫理的結果に関する膨大な量の研究論文を生み出した。概して、これらの間には重要な相違が存在するが、それらに共通するのは、広範なデータに関して数学的に構造化されたプロセスに基づいて、警察介入、保釈審査、量刑などの業務が行われることを推奨することである。それらは、現場の担当者が多種多様な規則・標準に従って下す判断に基づくより伝統的な法的決定手段と著しく異なる。新たに登場しつつある技術に直面している刑事司法において、どのような倫理的問題が生じているのか、検討する必要がある (Chi-

ao, 2019: 126)。

これらの問題は、公平性、説明責任、透明性の3つに分類することができる。第一に、新しい技術が基礎とするデータに偏りがあるとしたら、その技術が公平であると考えられるであろうか。第二に、技術に誤りが生じた場合、誰に非難の矛先を向けたいのであろうか。第三に、アルゴリズムの作用の方法、結論の導出過程を理解できないことは問題なのか（Chiao, 2019: 126; Dupont *et al.*, 2018）。

## 2 アルゴリズム使用最低基本三原則

### 1) 公平性

2016年5月、ProPublicaはCOMPASに関する詳細なレポートを発表し、CAMPASには人種の偏見があり不正確であるとした。その分析によると、暴力犯罪の予告点数は「著しく信頼できない」ことが証明されたばかりでなく、基本原則には公に人種が計算されていないにもかかわらず、重大な人種的不平等が含まれている。COMPASは、黒人被告人を実際よりも再犯を犯す危険性が高く、白人被告人を再犯率が低いと誤ったラベル付けをした（Kehl *et al.*, 2017: 28-29; Ferguson, 2017）。

偏見のリスクは、予測的警察取締りに使われるような潜在的に偏りのあるデータを使用するアルゴリズムによって悪化する。これらのアルゴリズムの間の相互作用は、O'Neilの主たる関心の一つである。警察は、「報告される」犯罪（暴力犯と財産犯）と「発見される」犯罪（職質で発見される少量の麻薬所持と違法活動への関与）の2つのタイプの犯罪に応答する。「発見される」犯罪に関して、圧倒的に貧困且つマイノリティーの住民は不釣り合いな量の警察活動に直面する傾向にある。その結果、「発見される」犯罪に関するデータは、貧困者とマイノリティーのコミュニティーは現実よりも高い割合でこれらの犯罪をするという偏った示唆をする傾向にある。そのような情報が再犯リスク計算にインプットされると、貧困者あるいはマイノリティーの被告人は将来犯罪をする大きな危険性があり、高いリスク点数が付与されるという誤った示唆がなされる（Kehl *et al.*, 2017: 29; O'Neil, 2016）。

これらのアルゴリズムの開発、とりわけ、入力と利用に関して、重要な問題と取り組まなければならない。リスク評価アルゴリズムは、潜在的に人種

を代用しあるいは歴史的に偏ったデータに依存する不必要な変数をできる限り排除すべきである。精度がある程度減じるとしても、最も問題のある変数の一部を取り除かなければならない (Kehl *et al.*, 2017: 34)。

さらに、ガイドラインのような手続的な予防手段によって、アルゴリズムが使われる刑事司法システムにおいて点数が適切に扱われ、偶然の影響を最小限にすることができる。これらの規則は、アルゴリズムに使われるデータを閲覧し、異議申し立てをする被告人の権利を規定しなければならない (Kehl *et al.*, 2017: 35)。

## 2) 説明責任

最大限の説明責任を促進するために、システムは特定の法管轄領域あるいは地理的地域に適切で、正確さと信頼性のために継続的にモニターされ評価されなければならない。いかなるツールも、最善の科学に基づき、問題のある変数が入り込むような潜在的な否定的影響を最小限にするよう測定され、完全無欠でなければならない (Kehl *et al.*, 2017: 33)。

有効性に関する研究も個々の法管轄毎に行われなければならない。国家レベルあるいは他の法管轄で検証されたツールは特定の場所では適切でないかもしれない。ローカルな政策決定者は、いかなるリスク評価システムを採用するかを決定する前に、そのような情報を持ち、決定の前提として有効性ほかに関する研究を行わなければならない。検証と有効性に関する研究は、一回限りでなく、繰り返し行われ、人口やツールに関する新しい情報の変化に基づいて、適切に変更されなければならない。政策立案者は、他の法管轄の同様の業務に携わる者と最善の実践について情報交換し、法管轄間の標準化の機会を探求し、偏見や信頼されないアルゴリズムから個人を守らなければならない (Kehl *et al.*, 2017: 33)。

外部研究や監査も重要である。外部研究者がアルゴリズムの偏見について評価し検証するデータとツールにアクセスできないならば、高度の透明性もほとんど意味を持たない。これらのツールも、現実的な現状改革を確実にするために、既存のメカニズムと比較して、厳格に評価されなければならない。研究プロジェクトは、ツールの機能の仕方、裁判官の現実の使用方法、最善かつ最適、大規模に配置する方法について、批判的な洞察を提示しなければならない (Kehl *et al.*, 2017: 33-34; O'Neil, 2016: 208-210)。

最後に、リスク評価アルゴリズムが関係する場合、これらのツールの採用ならびに使用に対する継続的監督管理は不干涉過程であってはならない。政策立案者は、全ての段階に関わり、営利目的企業、研究機関、非営利組織のいずれであっても、パートナーに難題を問い、ツールの開発・資料における前提と決定について説明し、正当性を証明させなければならない。政府は、リスク評価を民間企業に外部委託し、それらが個々の被告人にとって正確で公正な法的決定を導くであろうと仮定してはならない。政策立案者は、それらの使用を監督管理し、開発者とその雇用主がシステムの包括的目的について提携し、いかなる潜在的欠陥も認識するのを確実にすることに、積極的な役割を果し続けるべきである（Kehl *et al.*, 2017: 34）。

### 3) 透明性

法律ならびに技術専門家が強調する中心テーマの一つは、アルゴリズムの開発方法、すなわち、要因の評価方法、評価とアップデートの頻度を設計する際に作成される前提に関する高い透明性の必要性である。透明性だけで偏見の可能性を必ずしも減じられるのではないが、多数の理由から有用である。第一にそして最も重要なこととして、高い透明性は外部研究者による徹底的な検査を促進する。また、これらのシステムがどのように機能しているか、それらの適用における二律背反に関する一般人の理解を増加する。要因のインプットと評価に関するより多くの情報は、容認されないあるいは潜在的に容認されない要因の使用に基づく将来的な異議申し立てにとっても重要である（Kehl *et al.*, 2017: 32）。

アルゴリズムによる決定に異議申立てをする十分な機会を保障することを目的とする「技術的適正手続き」（*technological due process*）が提唱される。技術的適正手続きの中心価値は、透明性、精度、説明責任、参加、公平性である。個人データを収集し、収集データから点数を生み出し、決定者に点数を配分し、それらの点数を決定に使用する点数システムに対する国家的管理監督を強化するよう要求される。個人は、検査し、不正確なデータについて反論し、修正し、データの情報源・供給者を知る権利を持つべきである。データから点数を生み出したアルゴリズムは公開され、個々のプロセスが検査できるようにしなければならない。政策立案者は、点数が公正で正確で反復可能であることを保証する必要がある（Kehl *et al.*, 2017: 32; Citron, 2014）。

技術的適正手続きの背後にある主たるメカニズムは、アルゴリズムの規則と決定の間の相関関係を記録する監査軌跡の要求である。監査軌跡は、事実のマップと個々の決定に適用された規則を含む。販売者は、アルゴリズムのソースコードを一般に公開し、部外者がこれらのアルゴリズムを検証できるようにしなければならない（ソフトウェア開発者の標準的な実務）。検証は、人種、国籍、性的指向、ジェンダーに基づく問題のある分類のパターンを明らかにすることができる。データを公開することによって、研究者は点数システムを批評し、一般の価値観をシステムに含めることができる。イノベーションに脅威を与え、参加者がシステムをゲームすることを容易にするとし、情報公開に反対する者もいるが、そのような行為が発生する十分な証拠がないという事実により、そのような懸念は減じることができる。完全なる情報公開は理想的であるが、政策立案者は企業と協力して事例ごとに制限された形での適切な情報公開を決定することができる (Kehl *et al.*, 2017: 32-33; Citron *et al.*, 2014)。

著作権のあるリスク評価ソフトウェアを使用するか、研究者や非営利者と共同して特定の法管轄領域のための特別なツールを開発するかについての政府の決定について情報提供すべきである。COMPAS のような著作権のあるツールは、公的なツールと比べ、検証や監督管理に服すことは少ない。コンピュータ科学研究者は、Northpointe はその著作権のあるアルゴリズムの詳細を情報公開することを拒絶したがために、不公平である可能性の程度について十分に評価することができない。Northpointe は収益を守らなければならないということは理解できるが、営利目的企業にリスク評価ツールの開発を依存することに関する疑問が起こる。製品を保護し販売する民間企業の正当なビジネス利益と公的説明責任の必要性との間の緊張を解決するのは容易ではない (Kehl *et al.*, 2017: 33)。

### 3 小結

刑事司法システムにおいて、犯罪リスク予測、自動化のためにアルゴリズムが使われる機会がますます増加している。これらのアルゴリズム駆動ツールは、刑事司法システムにおける決定の正確性ならびに効率性を高め、人間の誤りと偏見のリスクを減じる潜在的可能性があるが、既存の偏見を強化・

悪化し、司法システムの重要な要素である公平性を損なう潜在的危険性もある。情報通信技術の発展によって、これらのアルゴリズム駆動ツールの使用はさらに拡大するであろうが、その際には民主主義の基本原則ならびに倫理規範に基づき、「公平性」「説明責任」「透明性」が不可欠かつ重要な基本原則として確立されなければならない。

## VI エピローグ

A Iやロボット工学は法執行にとって新しい概念であり、犯罪予防と統制を強化するために、既に、多数の法執行機関が積極的にA Iとロボット工学の採用を検討しているが、関与レベルはまちまちである。A Iとロボット技術を法執行において個別事件に実用することを構想し、組織内に専門機関を設立した国々もある。共通して言えるのは、各国の実情を踏まえ、経験を積み重ねるといふ姿勢である。

このように、法執行における潜在的可能性は多様であるが、これらの法執行に共通の目的は監視能力を高めることである。世界人権宣言（the Universal Declaration of Human Rights）、市民的及び政治的権利に関する国際規約（the International Covenant on Civil and Political Rights）、その他の国際的・地域的法律文書において承認されているように、いかなる監視の類型であっても、プライバシーに対する基本的人権の保障は考慮すべき最重要の問題である。法執行によるA Iとロボット工学の使用が社会に拡大し、市民生活と接触する機会が増大するに従い、これらの技術の使用が倫理的であることを保証することは、法執行にとってますます重要になっている。

プライバシーは安全の犠牲になりやすいがために、法執行はA Iとロボット工学の使用におけるプライバシーと倫理に関する重要なテストケースである。A Iとロボット工学の倫理的使用のために、法執行が規範を作成し、専門機関を設立すれば、他の領域も追随するであろう。また、A Iとロボット工学の使用が法執行において普及する前に、法執行はこれらの問題について議論する有利な立場にある。もし、この機会が無視され、「公平性」「説明責任」「透明性」なく、A Iとロボット工学が法執行において使用されるなら

ば、地域社会と市民の信頼性を失うリスクを負うことになるであろう。

最後に、民主主義原則が技術の世界でも構築されなければならない。この原則には、人権と人間の尊厳、自由と自己決定、多様性とマイノリティーの保護、権力の分割、チェックとバランス、参加の機会、透明性、公平性、正義、正当性、匿名性と投票の平等、誤用と暴露からの保護という意味でのプライバシー、干渉されない権利などが含まれる。将来において、法的、技術的に確立された原則は、「誰がどのデータをどのような目的で使用するのか、期間と価値を私たちが決める」ということである。個人データの使用、科学と政策のために作られた統計は透明性をもって個人に報告されるべきであろう。私たちのシステムを「多次元のリアルタイムのフィードバック・システム」にアップグレードしなければならない。そのような多次元的動機付けは、複雑システムを管理運営し、「自己組織化・自己規制システム」の実現を可能としなければならない。

#### 【参考文献】

- Bennett Moses, L., and Chan, J. (2018). Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and Society: An International Journal of Research and Policy*, 28 (7), 806-822.
- Berk, R. A., and Bleich, J. (2013). Statistical Procedures for Forecasting Criminal Behavior. *Criminology and Public Policy*, 12 (3), 513-544.
- Brayne, S. (2018). The Criminal Law and Law Enforcement Implications of Big Data. *Annual Review of Law and Social Science*, 14, 293-308.
- Brkan, M. (2017). Do Algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond. SSRN Scholarly Paper.
- Brucato, B. (2015). Policing Made Visible: Mobile Technologies and the Importance of Point of View. *Surveillance and Society*, 13, 455-473.
- Chan, J., and Bennett Moses, L. (2016). Is Big Data challenging criminology? *Theoretical Criminology*, 20 (1), 21-39.
- Chiao, V. (2019). Fairness, accountability and transparency: notes on

- algorithmic decision-making in criminal justice. *International Journal of Law in Context*, 15, 126–139.
- Christin, A., Rosenblat, A., and Boyd, D. (2015). Courts and predictive algorithms. *Data and Civil Rights: A New Era of Policing and Justice*. Data and Society.
- Cino, J.G. (2018). Deploying the Secret Police: The Use of Algorithms in the Criminal Justice System. *Georgia State University Law Review*, 34 (4), 1073–1102.
- Citron, D. K. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, 89 (1), 1–33.
- Douglas, T., Pugh, J., Singh, I., Savulescu, J., and Fazel, S. (2017). Risk assessment tools in criminal justice and forensic psychiatry: The need for better data. *European Psychiatry*, 42, 134–137.
- Dupon, B., Stevens, Y., Westermann, H., and Joyce, M. (2018). *Artificial Intelligence in the Context of Crime and Criminal Justice*. A Report for the Korean Institute of Criminology. Korean Institute of Criminology and International Centre for Comparative Criminology.
- Eubanks, V. (2017). *Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor*. New York: St. Martins Press.
- European Commission for the Efficiency of Justice (CEPEJ) (2018). *European Ethical Charter on the use of artificial intelligence in judicial systems and their environment*.
- Ferguson, A.G. (2017). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: New York University Press. (大槻敦子訳『監視大国アメリカ』原書房、2018年)
- Gates, K. A. (2011). *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. New York and London: New York University Press.
- Goel, S., Rao, J.M., and Shroff, R. (2016). Personalized Risk Assessment in the Criminal Justice System. *American Economic Review: Papers and Proceedings*, 106 (5), 119–123.

- Hamilton, M. (2015). Adventures in Risk: Predicting Violent and Sexual Recidivism in Sentencing Law. *Arizona State Law Journal*, 47 (1), 1–62.
- Hannah-Moffat, K. (2019). Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates. *Theoretical Criminology*, 23 (4), 453–470.
- Harcourt, B.E. (2007). *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: Chicago University Press.
- Harcourt, B.E. (2015). *Exposed. Desire and Disobedience in the Digital Age*. Cambridge: Harvard University Press.
- Katzenbach, C., and Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, 8 (4), 1–18.
- Kehl, D., Guo, P., and Kessler, S. (2017). *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*. Responsive Communities Initiative, Berkman Klein Center for Internet and Society, Harvard Law School.
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication and Society*, 20 (1), 14–29.
- Lash, S. (2007). Power after Hegemony: Cultural Studies in Mutation? *Theory, Culture and Society*, 24 (3), 55–78.
- Leclerc, B., and Cale, J. (eds.) (2020). *Big Data (Criminology at the Edge)*. London and New York: Routledge.
- Mehozay, Y., and Fisher, E. (2019). The epistemology of algorithmic risk assessment and the path towards a *non-penology penology*. *Punishment and Society: The International Journal of Penology*, 21(5), 523–541.
- O’Neil, C. (2016). *Weapon of Math Destruction: How Big Data increases Inequality and Threatens Democracy*. Penguin Books. (久保尚子訳『あなたを支配し、社会を破壊する A I ・ビッグデータの罠』インターシフト、2018 年)
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge and London: Harvard University Press.
- Peeters, R., and Schuilenburg, M. (2018). Machin justice: Governing security

- through the bureaucracy of algorithms. *Information Policy*, 23, 267–280.
- Perry, W.L., McInnis, B., Price, C.C., Smith, S.C., and Hollywood, J.S. (2013). *Predicting Policing: The Role of Crime Forecasting In Law Enforcing Operations*. Santa Monica: RAND Corporation.
- Rehavi, M. M., and Starr, S. B. (2014). Racial Disparity in Federal Criminal Sentences. *Journal of Political Economy*, 122 (6), 1320–1354.
- Saunders, J., Hunt, P., and Hollywood, J.S. (2016). Predictions put into practice: a quasi-experimental evaluation of Chicago’s predictive policing pilot. *Journal of Experimental Criminology*, 12 (3), 1–25.
- Schuilenburg, M. (2015). *The Securitization of Society: Crime, Risk, and Social Order*. New York: New York University Press.
- Smith, G.J.D., Bennett Moses, L., and Chan, J. (2017). The Challenges of Doing Criminology in the Big Data Era: Towards a Digital and Data-driven Approach. *The British Journal of Criminology*, 57 (2), 259–274.
- Takemura, N. (2020). *AI-Algorithm-Big Data, Predictive Criminal Justice and Hyper Crime/Social Control: Surveillance Capitalism after ‘Singularity’ and Prospects of Informational Civilization*. Background documentation from individual expert prepared by Noriyoshi Takemura. The 14<sup>th</sup> United Nations Congress on Crime Prevention and Criminal Justice, Kyoto, 20–27 April 2020. (forthcoming, the Congress was postponed because of COVID-19)
- The Committee of Experts on Internet Intermediaries (MSI-NET) (2017), *Study on the Human Rights Dimensions of Automated Data Processing Technique (in Particular Algorithms) and Possible Regulatory Implications*.
- Veale, M. (2019). *Algorithms in the Criminal Justice System*. A report by The Law Society Commission on the Use of Algorithms in the Justice System. The Law Society of England and Wales.
- 渡邊満久 (2018) 「AI 利用に伴うプロファイリングの諸問題」 *PwC Legal Japan News*, 1–5.
- Williams, M.L., Burnap, P., and Sloan, L. (2017). Crime Sensing With Big Data: The Affordance and Limitations of Using Open-source Communications to Estimate Crime Patterns. *The British Journal of Criminology*, 57 (2),

320-340.

Završnik, A. (ed.) (2018). *Big Data, Crime and Social Control*. London and New York: Routledge.

Završnik, A. (2020). Criminal justice, artificial intelligence systems, and human rights. *ERA Forum*, 20, 567-583.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30, 75-89.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

(たけむら・のりよし 桐蔭横浜大学法学部教授)