| | |
|---|---|
| Title | 予測的ポリシングと憲法 :<br>警察によるビッグデータ利用とデータマイニング |
| Sub Title | Predictive policing and the constitution |
| Author | 山本, 龍彦(Yamamoto, Tatsuhiko) |
| Publisher | 慶應義塾大学大学院法務研究科 |
| Publication year | 2015 |
| Jtitle | 慶應法学 (Keio law journal). No.31 (2015. 2) ,p.321- 345 |
| JaLC DOI | |
| Abstract | |
| Notes | 論説 |
| Genre | Departmental Bulletin Paper |
| URL | https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AA1203413X-20150227-0321 |

# 予測的ポリシングと憲法

## ――警察によるビッグデータ利用とデータマイニング――

山　本　龍　彦



## Ⅰ　はじめに

リチャーズとキング（Neil M. Richards and Jonathan H. King）によれば、過去半世紀にわたって我々の社会や我々の生活を変え続けてきた「情報革命」は、ここ数年で第三段階に入ったとされる[1)]。1970年代のインテル社によるマイクロプロセッサーの開発と、それによる計算能力（the power to compute）の向上によって画される第一段階、ネットワーク化と、それによる連結能力（the power to connect）の向上によって画される第二段階に続く、新たな段階である。それは、「データ」と、「予測能力（the power to predict）」の向上によって画される段階であるとされ、それ単独で「ビッグデータ革命（“Big Data” Revolution）」とも称される社会的転換期[2)]である。リチャーズとキングの言葉を借りれば、この「革命」によって「デート、ショッピング、医療、投票行動、法執行、テロ

1）Neil M. Richards and Jonathan H. King, *Big Data Ethics*, 49 Wake Forest L. REV. 393, 397（2014）.

2）*Id*. at 393.

対策、サイバーセキュリティを含む、ありとあらゆる人間活動および決定が、ビッグデータによる予測（predictions）によって影響を受け始めている」[3] という。

本稿は、ここで「影響を受け始めている」と名指しされる警察の法執行ないし犯罪予防の領域に焦点を当て、そこでのビッグデータ利用、とりわけ、これを前提としたデータマイニングおよびその適用――いわゆる予測的ポリシング（predictive policing）――の（憲）法的問題を考察することを目的とする。既に多くの論者が、「ビッグデータは、様々なかたちで、警察による捜査に次の劇的変化（dramatic change）をもたらしうるものである」[4] などと指摘しているが、本稿は、ビッグデータ利用の具体的実例を紹介しつつ、それがどのような点で警察活動に「劇的変化」をもらしうるのか、また、こうした実践がなぜ（憲）法的に問題になりうるのか――そこで生ずるのは、単に個人情報保護の問題だけなのか、など――を具体的に検討し、あくまで試論の範疇にとどまるものの、警察によるビッグデータ利用ないしデータマイニングの適用等に対する適切な法的統制のあり方を探求することとしたい。

## Ⅱ　何が問題なのか？

### 1　実　例

#### ⑴「エリア」に着目した予測

警察によるビッグデータ利用の実例として、まずは、カリフォルニア州サンタクルーズ郡の予測的ポリシングに関するプロジェクトを挙げることができる。2011 年 7 月に開始されたこのプロジェクトは、過去の厖大な犯罪データから一定のパターンを抽出・発見して、車両・住居への不法目的侵入（burglaries）および車両窃盗のような財産犯が発生しやすい場所と時間――“hot spots”――を予測するコンピュータ・プログラム（PredPol）を構築することを目的と

3）*Id*.

4）Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 Wash. L. Rev. 35, 37（2014）.

している[5]。この予測プログラムにより、有限の警察資源を効果的に配分すること――犯罪の発生しやすい場所・時間に警察官を重点的に配備することなど――が可能となり、予算削減と効果的なポリシングの両立が実現されたなどと指摘されている[6]。もちろん、これまでもCompStatのような犯罪追跡（crime-tracking）システムは存在し、実際に多くの警察によって利用されてきたが、サンタクルーズ警察のプログラム（ソフトウエア）は、犯罪が発生するごとにデータが蓄積され、プログラムの内容が日常的に更新・検証されることで、予測の精度が従前のものよりも飛躍的に高まったという（hot spotとして提示される範囲も絞り込まれ、現在では、約150メートル四方の正方形として地図上に表示される。毎日15のspotsが警察官に示され、警察官はウエブ上で常にその位置を確認できる）[7]。実際、同プログラムを導入した2011年7月と前年同月を比較すると、不法目的侵入の件数は27％（70件から51件）低下したと指摘されている[8]。また、コンピュータ・アルゴリズムの適用により、人間の経験や勘では思いつかないようなhot spotsが浮かび上がり、それが実際の逮捕や犯罪防止と結び付くことも少なくないとされる[9]。

他に、犯罪の発生する社会的文脈など、過去の犯罪の時間や場所以外の要素を考慮するアプローチも存在する。例えば、ニュージャージー州モリス郡の警察は、いわゆるリスク面分析（risk terrain analysis）[10]において、①過去の不法目的侵入、②財産犯の罪で最近逮捕された者の居住地、③主要幹線道路との距

5）Erica Goode, *Sending the Police Before There's a Crime*, N.Y. Times（Aug. 15, 2011）, http://www.nytimes.com/2011/08/16/us/16police.html.

6）*Id*.

7）Zach Friend, *Predictive Policing: Using Technology to Reduce Crime*, FBI Law Enforcement Bulletin（Apr. 9, 2013）, http://leb.fbi.gov/2013/april/predictive-policing-using-technology-to-reduce-crime.

8）Jennifer Bachner, Predictive Policing 25-26（2013）, available at http://www.businessofgovernment.org/sites/default/files/Predictive%20Policing.pdf.

9）日本における“hot spots”分析については、雨宮護＝島田貴仁「東京23区における住宅対象侵入窃盗犯の地理的分布の変化」都市計画論文集48巻1号（2013年）60頁以下等を参照。

離、④若者の地域的な凝集、⑤集合住宅およびホテルの位置情報を用いている[11]。郡警察によれば、このリスク面分析のために、暴力犯および財産犯の件数は「大きく低下」したとされる[12]。

また、ニューヨーク市警察は、マイクロソフト社と協力して、「領域認識システム（Domain Awareness System, DAS）」を開発し、採用している。このソフトウエアは、「市内に設置された約 3000 の監視カメラ、200 以上のナンバープレート自動読取装置、2000 以上の放射線センサー、警察保有のデータベースといった異なる源泉からの情報を常時取集、連結、分析するもの」[13] で、市内の「潜在的危険（potential threats）」（持主不明のバッグ等）を発見するために用いられている。ジョー（Elizabeth E. Joh）は、「このシステムは、個々の犯罪分析家にはわからないような方法で、人・物・場所の関連性を明らかにし、かつ、こうした情報に、警察がリアルタイムでアクセスすることを可能にするもの」[14] と説明している。報道によれば、市警察は、このソフトウエアによって、2013 年 11 月に開催されたニューヨークシティマラソン――同年 4 月のボストンマラソン後、テロの標的となりえた――のほぼ全ての場面を監視することができたとされる[15]。

### ⑵ 「個人」に着目した予測

さらに、フライトの安全の確保という観点から、空港において、いわゆる“No Fly List”該当者の同定が行われている。このリストの作成・準備段階で

---

10）*See e.g.*, Leslie W. Kennedy et al., *Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation*, 27 J. Quantitative Criminology 339, 342-343（2011）.

11）Joh, *supra* note 4, at 46.

12）Jeffrey S. Paul and Thomas M. Joiner, *Integration of Centralized Intelligence with Geographic Information System*, Geography & Pub. Safety, Oct. 2011, at 7.

13）Joh, *supra* note 4, at 48-49.

14）*Id*. at 49.

15）*See* Michael Schwirtz, *After Boston Bombings, New York Police Plan Tight Security at Marathon*, N. Y. Times（Nov. 1, 2013）, http://www.nytimes.com/2013/11/02/sports/video-surveillance-to-be-a-key-component-of-marathon-security.html.

データマイニングが行われているかははっきりしないが、「政府の願望を踏まえれば、この文脈で自動予測のプロセス（automated prediction processes）が適用されていたとしても、何ら驚かない」[16]とも指摘されている。“No Fly List”の手続とは異なるが、国土安全保障省（Department of Homeland Security, DHS）は、実際、入国管理に関して、データマイニングに基づく予測的モデリングを用いているとされる[17]。確かに、同省によるデータマイニング報告は、対人的標的化システム（Automated Targeting System-Persons, ATS-P）モジュールについて言及しており、国境を跨ごうとする者のリスクを見積もるために、政府保有の種々のデータベースが分析されることを示している[18]。

警察実務との関連でより注目すべきは、シカゴ警察（Chicago Police Department, CPD）による「便宜告知プログラム（Custom Notification Program）」であろう。ここでは、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」といったデータも含まれる）[19]――情報自由法（Freedom of Information Act, FOIA）に基づく情報公開請求が斥けられており、用いられるデータの詳細は明らかにされていない[20]――から、暴力犯の実行者と犠牲者を予測するアルゴリズムが開発され、その適用によってこれらの者を具体的にまとめたリスト――“heat list”と呼ばれる――が作成されている[21]（2014年2月の報道の時点では、400人を超える者がこのリストに掲載されているようである[22]）。警察は、heat listに掲載されている者を一人ひとり訪ね

16）Tal Z. Zarsky, *Transparent Predictions*, 2013 U. Ill. L. Rev. 1503, 1515（2013）.

17）*Id*.

18）*Id: see also* U.S. Dep't of Homeland Security Privacy Office, 2010 Data Mining Report to Congress 13（2010）, available at http://www.dhs.gov/xlibrary/assets/privacy/2010-dhs-data-mining-report.pdf.

19）Kristal Hawkins, *"Heat list' brings Minority Report-style police attention for likely offenders in Chicago,* Crime Library（Feb. 24, 2014）, http://www.crimelibrary.com/blog/2014/02/24/heat-list-brings-minority-report-style-police-attention-for-likely-offenders-in-chicago/index.html.

20）Aaron Rieke, David Robinson And Harlan Yu（Robinson + Yu）, Civil Rights, Big Data and Our Algorithmic Future 18（2014）, available at http://bigdata.fairness.io/wp-content/uploads/2014/11/Civil_Rights_Big_Data_and_Our_Algorithmic-Future_v1.1.pdf

（手紙を送る場合もある）、将来犯罪に手を染めた場合の結果について警告するとともに、その者が受けられる社会的サービス（職業訓練、住宅供給等）を告知するものとされている[23]。CPD は、こうした努力によって、犯罪が抑止されてきていると主張し、シカゴ市市長（Rahm Emanuel）も、2014 年 2 月報道の時点で、これまでなされた 60 の訪問ないし介入により告知を受けた者は、いずれも新たな重罪に手を染めていないと主張している[24]。このような効果が謳われる反面で、「heat list は、『犯罪前（pre-crime）』の法執行を恐ろしくも現実化する差別的なツールである」[25] との主張があるように、厳しい批判も寄せられている。実際、重罪犯等にかかわったことのない者がリストに掲載され、それがために突然警察の訪問を受けるケースもあり、リスト掲載者の多くが、その掲載の事実について当惑を感じていると指摘される[26]。このような予測的ポリシングの負の側面については後に詳しく検討したい。

未だ実験段階とされているものの、既に複数の専門家ないし技術者が、ある者が将来重罪を犯すかどうかを一定の精度で予測できるソフトウエアを開発したと宣言している。例えば、バックグラウンドチェックを業務内容とする Intelius 社にプライバシー・オフィサーとして勤務していたアドラー（Jim Adler）は、同社保有の厖大なデータ――重罪犯歴、軽犯罪歴、交通違反歴、ジェンダー、目・肌の色、タトゥーの有無、等々――に基づき、ある者が重罪を犯すかどうかを「合理的な正確性（reasonable accuracy）」をもって決定するアルゴリズムを開発したと述べている[27]。報道によると、アドラーのプログラ

21）Jeremy Gorner, *Chicago police use 'heat list' as strategy to prevent violence*, Chicago Tribune（Aug. 21, 2013）, http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821_1_chicago-police-commander-andrew-papachristos-heat-list.

22）Hawkins, *supra* note 19.

23）Rieke, *supra* note 20, at 18. なお、CPD は、「警告」よりも、社会サービスの告知という点を強調して、「便宜告知プログラム」と呼んでいるように思われる。

24）*Id*.

25）Hawkins, *supra* note 19.

26）*Id*.

ムは、最もアグレッシブな設定の下では、1980年代以降、ケンタッキー州の裁判記録上重罪を犯したとされた全ての者（5万1246人）を正しく同定した一方、2220人の非重罪犯罪者を誤って同定したという[28]。ただ、より寛大な設定の下では、3万7842人の重罪犯罪者を正しく同定したのに対して、誤って同定した非重罪犯罪者は152人にとどまったとされる[29]。アドラーは、この偽陽性（false positive）がさらに減じられれば、警察による実装の可能性は非常に高いと指摘するが、当然慎重論も存在している。アドラー自身は、このような「個人」標的的な予測的ポリシングが正しい行いか否かを考えるのは専門家や技術者――彼は「私のようなオタクたち（geeks）」と呼ぶが――の仕事ではないと述べるが[30]、少なくともそれは法学者の仕事ではあるであろう（本稿の目的の1つは、この「仕事」の起点をなすことである）。

他にも、ペンシルバニア大学で統計学・犯罪学を教えるバーク（Richard A. Berk）らは、仮釈放中の者が殺人に関与するかどうかを予測するソフトウエアを開発したと主張している[31]。バークらによれば、6万人以上のデータを用いて作成されたこのソフトウエアは、「仮釈放した場合の殺人への関与を75％以上の確率で予測できる」[32]という（これは同時に、「4回に1回は判断を誤る」ということを意味するが、こうした指摘を踏まえた法的分析は、3で行うこととする）。さらに、メリーランド州では、バークの協力の下、児童虐待を行う家族を予測

27）Jordan Robertson, *How Big Data Could Help Identify the Next Felon――Or Blame the Wrong Guy*, Bloomberg（Aug. 15, 2013）, http://www.bloomberg.com/news/2013-08-14/how-big-data-could-help-identify-the-next-felon-or-blame-the-wrong-guy.html.

28）*Id.*

29）*Id.*

30）*Id.*

31）Richard A. Berk et al., *Forecasting Murder within a Population of Probationers and Parolees: A High Stakes Application of Statistical Learning*, Journal of the Royal, Statistical Society（Series A）, 172（2009）at 971-1008. この研究については、ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ（斎藤栄一郎訳）『ビッグデータの正体』（講談社、2013年）242頁に紹介がある。

32）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注31）242頁参照。

するプログラムが開発中であるとの報道もなされている[33]。

## 2 焦 点

### ⑴「ビッグデータ」と「データマイニング」

以上、警察によるビッグデータ利用ないしデータマイニングの実例を簡単に紹介してきた。以下では、その憲法上の問題について考察することとするが、その前に、「ビッグデータ」や「データマイニング」といった言葉がもつ意味について若干の検討を加え、考察の射程――警察による情報処理のなかで、果たしてどのような行為を問題とすべきなのか――を明確にしておきたい。

まず、「ビッグデータ」であるが、これを主に技術的観点から、「伝統的なデータベースシステムのもつ処理能力を超えたデータ」[34]とか、「量、速度、種類がいずれも圧倒的な情報資源で、高度な洞察と意思決定のために効率的で革新的な情報処理技術を必要とする」データ[35]と定義する見解がある。ただ、近年は、技術的観点に社会的観点を加味した定義が有力に説かれている。例えば、マイヤー＝ショーンベルガーとクキエ（Viktor Mayer-Schönberger and Kenneth Cukier）は、「小規模ではなしえないことを大きな規模で実行し、新たな知の抽出や価値の創出によって、市場、組織、さらには市民と政府との関係などを変える」データ[36]と定義している。それが与える広範な社会的インパクトを踏まえれば、データの量などに加えて、そこから得られる洞察ないし推論・予測

33）*Predictive policing: Don't even think about it*, The Economist（Jul. 20, 2013）, http://www.economist.com/news/briefing/21582042-it-getting-easier-foresee-wrongdoing-and-spot-likely-wrongdoers-dont-even-think-about-it. さらに、Liane Colonna, *A Taxonomy and Classification of Data Mining*, 16 SMU SCI. & TECH. L. REV. 309, 359（2013）.

34）この定義は、さらにこう続ける。「データがあまりに大きく、あまりに早く動くために、あなたの限定的なデータベース構造には適合しないものである」、と。Edd Dumbill, *What Is Big Data?: An introduction to the big data landscape*, O'Reilly Radar（Jan. 11, 2012）, http://radar.oreilly.com/2012/01/what-is-big-data.html.

35）IT Glossary: Big Data, Gartner, http://www.gartner.com/it-glossary/big-data/（last visited Dec. 1, 2014）.

36）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）18 頁。

の社会的意味を考慮した捉え方をしておくべきであろう[37]。

次に、「データマイニング」であるが、これは、文字どおり、データ（data）を・掘り当てる（mining）ことを意味しない。基本的には、「データから・知識を・掘り当てること（knowledge mining from data）」[38]を意味するのであり、ハンとカンバー（Jiawei Han and Micheline Kamber）の指摘によれば、「データマイニング」とは本来「誤称（misnomer）」である[39]。ただ、ここでは、その通俗性から「データマイニング」という言葉を使いつつ、その意味を上記のように「データから知識を掘り当てること」と理解しておきたい。

「データマイニング」の定義については、現在も複数のものが存在している[40]。「データを拷問にかけて何かを白状させること」[41]といったユーモラスなものもあるが、「データのなかに隠れたパターンと微妙な関係性を明らかにし、将来の結果予測を可能にする規則性を推論するために、統計学的分析やモデリングのようなデータベース技術と技巧を適用すること」[42]というアメリカの会計検査院（Government Accountability Office, GAO）の定義や、「厖大なデータセットのなかから、これまでに知られていない、有効なパターンないし関係性を発見するために、統計学的モデル、数学的アルゴリズム、機械学習法のような洗練されたデータ分析法を用いること」[43]という議会調査局（Congressional

37）社会的意味については、「特集　ポスト・ビッグデータと統計学の時代」現代思想 2014 年 6 月号所収の諸論稿を参照。ビッグデータの性格について憲法的な考察を加えたものとしては、宮下紘「ビッグデータの活用とプライバシー保護」法学セミナー 707 号（2013 年）8 頁以下参照。

38）Jiawei Han & Micheline Kamber, Data Mining: Concepts and Techniques 6（3rd ed. 2011）.

39）*Id*.

40）詳細については、Colonna, *supra* note 33, at 310-312.

41）Jeff Jonas, *What is Data Mining? Depends Who You Ask...*, Jeff Jonas（Sept. 8, 2006）, http://jeffjonas.typepad.com/jeff_jonas/2006/09/what_is_data_mi.html.

42）U.S. Gen. Acct. Off., GAO-04-548, Data Mining: Federal Efforts Cover a Wide Range of Uses（2004）.

43）Jeffrey W. Seifert, Crs Report for Congress: Data Mining: An Overview: December 16, 2004 - Rl31798.

Research Service, CRS）の定義が一般的であろう。

### ⑵ 議論の焦点――分析／適用／介入

ここで注意を要するのは、データマイニングそれ自体が、個人の権利利益と直接関係するわけではない、ということである。上記の一般的な定義から窺い知れるように、データマイニングは、厖大なデータセット（いわばビッグデータ）のなかから統計的に有意なパターンないし関係性を抽出・発見しようとするものであり、使用する個々のデータが誰のものかについて、基本的に関心がないからである。データマイニングそれ自体から明らかになるのは、＜無香料ローションの購入歴＋特定サプリメントの購入歴＋大きめのバッグの購入歴⇒該当者は妊娠している可能性が高い＞[44]とか、＜男性である＋ヘーゼル色の瞳をもつ＋交通違反を超えた軽犯罪歴をもつ＋タトゥーを入れている⇒該当者は重罪を犯す可能性が高い＞[45]というパターンないし関係性であり、＜誰が妊娠しているか＞、あるいは＜誰が重罪を犯しやすいか＞ではない。要するに、使用されるビッグデータが匿名化されている限り[46]、データマイニングが特定個人と関係を結ぶのは、そこで抽出・発見されたパターン等が、個人識別情報を含むデータベースに「適用（apply）」される段階である（以下、この適用行為を「＜適用＞」と表記する）。ビッグデータから掘り当てられた上記のようなパターンがデータベースに＜適用＞されてはじめて“特定の誰か”の傾向や属性等が割り出されるのである。

こうした検討から明らかになるのは、個人の権利利益の侵害を基本的に検討しなくてもよい――法的分析の直接の対象としなくてもよい――「データマイニング」も存在するということである。例えば、前記１⑴で紹介した「エリア」着目的な予測は、そこで使用されるビッグデータが匿名化されている限り、

44）例えば、タオル５パック、敏感肌用の洗剤、ゆるいジーンズ、DHA を含むビタミン剤、大量の保湿剤を購入している者が妊娠している確率は 96％であるという。チャールズ・デュヒッグ（渡会圭子訳）『習慣の力』（講談社、2013 年）268-269 頁参照。

45）Robertson, *supra* note 27.

46）匿名化の困難性については、宇賀克也＝宍戸常寿＝森亮二「鼎談　パーソナルデータの保護と利活用へ向けて」ジュリスト 1472 号（2014 年）iv 頁等参照。

個人の権利利益と直接かかわるものではない。そこで標的にされているのは、あくまでも「場所」や「時間」であり、特定の「個人」ではないからである。他方、前記1⑵で紹介したデータマイニングの実例は、そうはいえない。例えば、CPD の heat list は、伝統的な地理的な犯罪マッピングを超えて、警察が早期に介入すべき「個人」を特定しようとするものである（エリアから人へ）。先述のとおり、データマイニングそれ自体、すなわち、データセットから従前は知られていないパターンないし関係性を抽出・発見すること自体は、個人の権利利益と直接関係するものではないが、ここでは、かかるパターン等が個人に――厳密には個人識別情報を含むデータベースに――＜適用＞され、「個人」の傾向や属性が判断されているのである。

以上のように考えると、“警察によるビッグデータ利用やデータマイニングは法的に問題だ”と抽象的かつ包括的に主張することは、議論の発展に何ら貢献しないということになる。上記の考察からも、主に法的な光を当てるべきは、データマイニングそれ自体というより、その結果――パターンや関係性――を「個人」に＜適用＞すること、すなわち、発見されたパターン等に当てはまる誰かを探索することであろう。もちろん、この結論にもいくつかの問題を指摘することができる。例えば、本稿は、データマイニングに関連するプロセスを、さしあたり、①パターン等の分析・発見の段階（データマイニングそれ自体）――匿名のフェーズ――と、②かかるパターン等の＜適用＞の段階――顕名のフェーズ――に分類したが、プログラム上、あるいはシステム上、両者を分離できないという指摘もありえよう。“抽出しながら適用する”、という連続的な作業工程である。このように、①と②を分離できない・しないのであれば、両者を一体的なものとして捉え、これを丸ごと――すなわちデータマイニングも含めて――法的検討の対象とせざるをえないであろう。かくして、法的な観点からは、両段階にあえて「壁」を設けることが望まれる。

また、そもそも＜適用＞すら法的検討の対象とすべきではないとする指摘もありえよう。実際のところ、個人への＜適用＞段階においても、警察官と対象者との物理的な接触はない。＜適用＞結果に基づいて、警察官が実際に行動に

移す段階――③行動・介入の段階――に至ってはじめて、対象者は自らが警察の関心事になっていることを具体的に知り、また何らかの現実的な不利益が生じることになるからである。そうすると、法的に統制すべきなのは、＜適用＞結果に基づく警察官の具体的な介入の段階であり、それ以前の①・②段階ではない、ともいえそうである。これは、決して軽視することのできない洞察である。以下では、＜適用＞がいかなる憲法上の権利・自由・価値と抵触するのかを具体的に検討することで、こうした指摘に応答してみたい。

## 3　＜適用＞の侵害的性格

データマイニングにより得られた結果（特定のパターンないし関係性）を、「個人」（個人識別情報を含むデータベース）に＜適用＞すること、別言すれば、データマイニング結果を利用して、特定個人の性質や傾向を割り出すことが、いかなる憲法上の権利・自由・価値を侵害することになるのであろうか。

### ⑴　プライバシーの権利

第1に考えられるのは、やはりプライバシー権（憲法13条）であろう。私生活ないし私事を密かに覗き見る行為がプライバシー権の侵害を構成するという点に異論がないとすれば[47]、データマイニング結果の＜適用＞は、既に保有している断片的個人情報からは明らかにならなかった情報主体の私事を新たに知ろうとする行為――データ媒介的覗き見――として、プライバシー権を侵害するものと考えることができる。例えば、妊娠しているか否かは（とくに外見的な特徴の表れない妊娠初期においては）「私事」に属する事柄であり、通常、本人の同意なくこれを知るには、法的に高いハードルが課されることになろう。しかし、データマイニング結果の＜適用＞は、一見すると相関性の認められない個別の商品の購入歴等から、本人とのいかなる接触もなく、その「事実」を知ることができるのである。もちろん、“ここでの結果は、その情報主体は妊娠しているかもしれない、という可能性に過ぎない”との反論もありえよう。パターンの＜適用＞により表出した結果は、ある者が妊娠しているという「事

47）東京地判昭和39年9月28日下民集15巻9号2317頁（「宴のあと」事件）参照。

実」ないし「真実」そのものでなく、分析者の創造的な知的作業に基づく「評価」である、との反論である。

しかし、プライバシー権に関するリーディング・ケースである「宴のあと」事件判決[48]が、「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」（傍点山本）と述べていることからすれば（「プライバシーの侵害は多くの場合、虚実がないまぜにされ、それが真実であるかのように受け取られることによって発生することが予想される」ともいう）、一定の精度が担保されたアルゴリズムによって導かれた結果は、「真実らしく受け取られる」情報であると解され、その限りにおいて、＜適用＞行為は「データ媒介的覗き見」として、プライバシー権の侵害を構成するものと考えることができよう[49]。

### ⑵ 個人の尊重原理、思想・良心の自由

第2に考えられるのは、憲法13条が謳う個人の尊重原理ないし個人の尊厳原理である。上記⑴で検討した、妊娠しているかどうかを推測するアルゴリズムの＜適用＞は、対象者の「過去」ないし「現在」の事実、あるいは対象者自身が知っている（自覚している）事実を割り出そうとするものであった。他方、前記Ⅱ1⑵で紹介したheat listのように、ある者が重罪を犯すかどうかを予測するアルゴリズムの＜適用＞は、対象者の「未来」に属する事柄、あるいは対

48）前掲注47）参照。

49）「宴のあと」事件判決は、「一般の人」にとって「真実らしく受け取られる」かどうかを基準に、そのプライバシー侵害性を判断している。この点、警察実務において＜適用＞結果を一次的に取扱うのは「一般の人」ではなく「専門家」であり、“データマイニング結果の＜適用＞によって得られた情報は、「事実」ないし「真実」そのものではなく、あくまでも統計的に導出された「推測」に過ぎない”と冷静に判断される可能性がある。そうすると、当該情報は、情報主体の「個人情報」というより、専門家の知的営為に基づく「評価」（知識）であると解することもできよう。しかし、このように解するためには、＜適用＞結果を取扱う警察官に対して、データマイニング等に関するリテラシー教育が徹底されること、その結果がかかる教育を受けていない「一般の人」の手に渡らないことなどの条件を満たすことが必要となろう。

象者すら知らない（自覚していない）事柄をまさに「予測」しようとするものである。このように、本人すら知らないその者の傾向、無意識的な欲動（心的表象）を覗き見て、将来の行動を「予測」することは、過去や現在の事実を割り出すこととは次元の異なる問題を提起するように思われる。例えば、リチャーズとキングは、このような「個人」の予測は、その人らしさ、つまりアイデンティティそのものを危うくすると主張している。「ビッグデータに基づく分析は、自分がそう決める前に、私とは何者であるかを調整し、決定（determinate）しさえする制度的監視を可能にする」[50] というのがその理由である。

また、マイヤー＝ショーンベルガーとクキエは、スモールデータに基づく古典的なプロファイリングが、集団ベースの分析を中心とすることで、「特定集団への差別につながる」[51] のに対し、ビッグデータに基づく予測は、「個人」ベースのより緻密な分析を中心とするため（「アラブ系の名前の客が片道キップでファーストクラスに乗ろうとしても、他のデータからテロリストの可能性がきわめて低い場合、空港の保安検査で別室に呼ばれることはなくなる」[52]）、差別的な要素や偏見を排除できるという[53]。しかし彼らは、そうであるがゆえに、つまり「個人」のより正確な把握が可能になるがゆえに、実際の＜私＝X＞と、予測アルゴリズムによって同定された＜私＝X′＞とのギャップが縮減し――＜私＝X＞が、＜私＝X′＞に「先回り」[54] され、追い込まれる――、＜私＝X＞は「確率という名の牢獄」に放り込まれることになると指摘している[55]。

50）Richards and King, *supra* note 1, at 422.

51）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）240 頁。

52）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）241 頁。

53）*See e.g.*, Joh, *supra* note 4, at 57-59.

54）「先回りされる個人」という表現につき、宍戸常寿「通信の秘密に関する覚書」高橋和之先生古希記念『現代立憲主義の諸相（下）』（有斐閣、2013 年）520-521 頁参照。

55）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）244 頁。「余剰」（ギャップ）縮減に基づく訂正困難性については、山本龍彦「遺伝子プライバシー論」憲法理論研究会編『憲法学の最先端』（敬文堂、2009 年）41 頁参照。

このような見解を踏まえると、データマイニング結果を用いた個人の傾向把握ないし行動予測は、憲法の基底的原理である個人の尊重ないし個人の尊厳と抵触しうるように思われる[56]。リチャーズとキングが指摘するように、それは、本人が自らの自由意思に基づいて選択し、行動する前に、その個人を統計的に「判断」しているからである。これは、自律的な個人としての主体性を否定するものともいえる[57]。個人の自律的・主体的な判断・選択や実際の努力を捨象して、もっぱら統計的な予測――厄介なことにその精度は低くないのであるが――から、その個人がいかなる者であるかを決めつけているためである（「差別」が、集団への偏見や固定観念によって非科学的に個人を決めつけるものであるのに対して、アルゴリズムによる予測は、ビッグデータの力によって、高度に科学的に、また統計的に個人を決めつけるものであるといえる。それらはともに「個人の尊厳」と関連しているといえよう）。リチャーズとキングが、ビッグデータによる個人の予測は、「私が何者であるかを決定する基本的権利（the fundamental right to define who I am）」を危うくする[58]と指摘するように、予測のための＜適用＞は、憲法の基礎にある個人の尊重原理と抵触するものといえるであろう[59]。

また、こうした予測は、憲法19条によって保障される思想・良心の自由と

56）マイヤー＝ショーンベルガーとクキエも、「これから我々個人にとって怖いのは『プライバシー』よりも『確率』となる」と述べ、「おそらくビッグデータの時代には、個人の尊厳を守る新たなルールが必要になる」と指摘している。ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注31）32頁。

57）個人の自律性（自律的個人）を基底とした重厚かつ緻密な憲法論を展開するものとして、佐藤幸治『日本国憲法論』（成文堂、2011年）がある（とくに172-177頁以下参照）。

58）Richards and King, *supra* note 1, at 423.

59）この点で、行動予測機能の付いた監視カメラの法的位置付けが問題となる。これも、ある種の「パターン」が映像の被写体に「適用」されているわけであるが、これが純粋に「個人」を対象にしたものか、それとも実質的に「エリア」を対象にしたものかについては、慎重な議論が必要である。被写体の――標準パターンからは――逸脱した行動により、何らかのアラームが作動するとしても、それが個人を標的にしているのかエリアを標的にしているのか、俄かには判断できない。当該システムが、被写体「個人」に関心を有するものか、「エリア」に関心を有するものか、といった検討が必要である。

も抵触しうる。ある者の過去・現在ではなく、未来の行動を予測するという行為は、行動の前段階である内心領域の動向を覗き見る行為ともいえるからである。例えば、ある者が重罪を犯すかもしれないという予測は、本人すら気づいていない、犯罪に対する無意識的な欲求を可視化しようとする行為である（もちろん、その欲求が本人に意識されたとき、これを本人が理性的に抑え込むことはありうるわけであるが）。無論、思想・良心の自由の保障範囲については種々の議論があるが、仮にこれを広く捉えれば、特定の行動へとつながる内心領域の動向（欲動等）や傾向を――ビッグデータを用いて高い精度をもって――割り出そうとする＜適用＞行為は、同自由を侵害するものと解する余地はあるように思われる。

もちろん、前記(1)でも触れたように、データマイニング結果の＜適用＞によって導出される個人の予測は、あくまでも統計的な推測であって、それによって個人を決めつけるものでも、個人の内心の動きを完全に言い当てるものでもない――したがって個人の尊重原理に抵触するものでも、思想・良心の自由を侵害するものでもない――と主張することは可能である。しかし、人間が、合理的正確性（reasonable accuracy）が検証された予測アルゴリズムの＜適用＞結果に対し、どこまで懐疑的・批判的になれるのか、疑問がないわけではない。非科学的な固定観念からも逃れられなかった人間が、科学的な確率から自由でいられることはできるのか（「人間は意外に“データの独裁”に支配されやすい」[60]との指摘は軽視できない）。このような観点に立つと、現実の＜私＝X＞が、データに基づく＜私＝X′＞に取って代えられること（＜私＝X＞が尊重されないこと）、あるいは、内心領域に、将来特定の行動へと結び付く欲動が存在していると受け取られること[61]はありうるように思われる[62]。

### (3) もっともな批判

以上みてきたところによれば、データマイニングに関連する＜分析→適用→

60）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）248 頁。

61）「他者」による認識を重要視した 19 条解釈論については、堀口悟郎「人格と虚像」慶應法学 30 号（2014 年）37 頁以下参照。

介入＞というプロセス中の＜適用＞は、それ自体、憲法上の権利・自由・価値を侵害する可能性がある。しかし、先述のとおり、問題は＜適用＞結果に基づいて警察官が実際にどのように行動するか（介入段階）であり、そこではじめて権利利益と具体的に衝突するという考えもありえよう。そして、こうした見解においては、＜適用＞結果を鵜呑みにしたような数字崇拝的な行動は、第三者からの統制が及ぶため、警察によるビッグデータ利用ないしデータマイニングは、結局、新たな法的問題を生じさせるものではない、と考えるかもしれない。例えば、警察官が、“Xが重罪を犯す可能性は75％である”、という＜適用＞結果のみに基づき、Xに対する強制処分を行おうとする場合、当然、裁判所による審査が入り、令状が発布されないということはありうる。このように、強制処分の段階で司法のコントロールが十分に及ぶとすると、＜適用＞がXの権利利益を直接侵害することはないともいえそうである[63]。

しかし、裁判所の統制が及ぶのは、＜適用＞に基づく警察官の行動が可視化する段階であるということには注意が必要である。現状の法制度を前提にすると、具体的介入に至る前の標的化（例えば、本人に気付かれない監視の段階）に対しては、裁判所による実質的な統制が及ばないということになるからである。この帰結は、かかる不可視的な標的化でも、個人の行動を萎縮させることがありうる[64]ということを踏まえると、問題があるように思われる。例えば、多数派からは「逸脱（deviant）」とみられるような特徴を複数もつ（が、犯罪にか

62）なお、前掲注48）で触れたように、「個人」の予測結果を扱うのが「一般の人」ではなく「専門家」であれば、かかる結果を冷静に受けとめ、＜私＝X＞と＜私＝X′＞とのギャップを認識し続けることは可能かもしれない。そうすると、データマイニング結果の＜適用＞は、個人の権利利益や個人の尊重原理と抵触するものというより、単なるインテリジェンス活動と捉えることができるのかもしれない。しかし、こう考えるためには、前掲注48）で挙げたような諸条件を満たすことが必要がある。

63）確かに、＜適用＞によって本人が直接傷つくことはない。しかし、プライバシー侵害行為と実害がズレることは少なくないように思われる。例えば、AがBの寝室に盗聴器を仕掛けたとしよう。Bがこれによりショックを受け、深く傷つくのは、Aによって盗聴器を仕掛けられたことを知った後であろうが、プライバシーの侵害は、盗聴器を仕掛けるというAの行為それ自体によって、既に生じているといえる。

かわろうなどとは思っていない）者が、＜適用＞によって犯罪者となるリスクを高く見積もられ、警察による重点的監視の対象となっているのではないかという意識をもつこと、それによって行動を差し控えること、あるいは行動を「標準化」することは十分にありうる[65]。このように考えると、具体的な介入以前の＜適用＞行為自体に法的な問題を見出し、これを適切に統制していく必要は、やはり高いように思われる（そうでないと、標的化ないし不可視的な監視は常態化しうる）。

＊　このことは、データに駆り立てられた具体的介入に対する司法的統制の重要性を否定するものではない。裁判所は、強制処分を認めるに当たり、その必要性・合理性を慎重に考慮するのはもちろん、原則として、＜適用＞結果、すなわち統計的な推測以外の根拠を求めるべきであり、＜適用＞結果にウエイトが置かれる場合であっても、後述するようなアルゴリズミスト（専門家）の助言等を得つつ、データの信頼性やデータ処理プロセスのインテグリティを吟味すべきである。ただ、それには、処理プロセスやアルゴリズムなどが裁判所に開示（disclose）されることが必要となる。この手続をどのように制度化するかも、今後の検討課題となる。

## Ⅲ　どのように統制すべきか？

これまでの考察によれば、データマイニングにより抽出・発見されたパターンや関係性を、①情報主体の過去・現在の私事を覗き見るために、あるいは②情報主体の未来の行動を予測するために（行動に至る前の内心の動向を覗き見るために）＜適用＞することは、憲法上の権利・自由・価値を侵害する可能性がある。とはいえ、憲法上の権利・自由・価値を侵害する国家の行為が全て違憲となり、無効となるわけではない。国家の側がその侵害を「正当化」できれば、それらの行為は憲法に適合するものとして維持される。そうすると、次に議論すべきは、これらの行為が憲法上正当化されるか、である。

### 1　個別的同意――その困難性

もちろん、警察が＜適用＞行為を行うことを事前に情報主体に告知し、同意

64）*See e.g.*, Neil M. Richards, *The Dangers of Surveillance*, 126 Harv. L. Rev. 1934, 1948-1952 (2013).

を得さえすれば、そもそも権利・自由の「侵害」は認められない（個人の尊重原理という客観法への侵害については、その限りではないが）。しかし、警察によるデータマイニング結果の＜適用＞を考えたとき、この個別的同意というアプローチは、以下の2つの点で限界がある。1つは、同意を得るためには、警察が情報主体に対し、＜適用＞に関する情報を事前に告知することが必要となるが、それが捜査上の秘密の開示等につながり、犯人の逃亡や罪証の隠滅等、公共の安全と秩序の維持に重大な支障を及ぼすおそれがある、ということである。また、事前の告知をなしえたとしても、実際の犯人は同意を拒否するかもしれないし、犯人ではない者は（同意を拒否すると疑われるために）実際上同意を拒否できないかもしれない（同意の事実上の強制）。

もう1つは、現実的に困難である、ということである。本稿は、便宜上、データマイニングにより抽出・発見されたパターン等を「個人」に＜適用＞するという言い方をしてきたが、厳密には、それは個人識別情報を含む「データベース」に対して＜適用＞することを意味する。「データベース」にパターン等を当てはめて、そこからリスクの高い者を絞り込むというのが、おそらく一般的な＜適用＞ということになろう。そうすると、＜適用＞は、最終的には個人を特定するものであるとしても、現実には、データベースに含まれる全ての者を対象に行われることになる。データマイニングの結果は、「ビッグデータ」から抽出され、「ビッグデータ」へと適用されるのである。このように考えると、＜適用＞の同意は、対象となるデータベースに含まれる全ての者から取得しなければならないということになるが、これは現実的にみて困難といわざるをえない[66]（なお、仮に＜適用＞行為を「強制処分」と考え、裁判所による令状を要求するとしても、令状に記載すべき対象者をどうするのか、裁判官はデータベース登録者の全てを対象とする包括的・網羅的な捜査を審査し、許容することができるのかなどが問題とされうる）。

---

65）*See id*. at 1948-1949. リチャーズは、これを「監視の標準化効果（normalizing effect）」と呼んでいる。

66）*See* Zarsky, *supra* note 16, at 1543.

結局、厖大な量のデータを扱うビッグデータ時代には、「プライバシーは死んだ（Privacy is dead）」との宣言は言い過ぎであるとしても、本人による自己情報のコントロール（個別の同意）は現実として「死んでいる」ことになる[67]。ビッグデータ時代にあって、人間が1つ1つの情報の動きを把握し、それぞれに本人から同意を得ること（本人の実質的・事実的なコントロールを認めること）は、実際上不可能といえるからである。しかし、このことは、理念としての「自己情報コントロール」、理念としての「同意」を否定することにはならない。本人のコントロールや同意が実際上不可能であることによって、＜適用＞行為の権利侵害性が打ち消されるわけではないのである。そうであるならば、＜適用＞行為の侵害性を認め、その憲法的正当化の必要性を認めたうえで、理念としての「コントロール」や「同意」を別のかたちで実現することが要求されるように思われる。すなわち、集合的同意としての法律の制定と、本人のコントロールに代わってその適切な運用を担保するための仕組み・構造の組み込みである。以下、この点についてやや詳しくみていきたい。

## 2　集合的同意＋構造（＋理由）

以上述べてきた理由から、＜適用＞につき個別の同意取得（本人の実際のコントロール）が困難であるならば、＜適用＞の侵害的性格を認めたうえで、その憲法的正当化を図らざるをえない。この点で、まず考えられるのは、集合的同意ともいうべき「法律」によってこれを民主的に是認することである。ただ、この法律の制定は、個別的同意の代替という消極的な意味付けを超えた意義をもちうる[68]。それは、＜適用＞が、個人の権利・自由の侵害を超えたインパクトをもつことと関係している。例えば、とりわけ個人の行動予測を目的とする＜適用＞は、警察による事前的・予防的介入を広く許すこととなり、国家の

67）*See* Richards and King, *supra* note 1, at 365-366, 409-413.

68）ここでは、強制処分法定主義と法律の留保論との関係に関する議論には踏み込まない。問題の所在については、亀井源太郎＝宍戸常寿＝曽我部真裕＝山本龍彦「［座談会］憲法と刑事法の交錯（前篇）」法律時報86巻5号（2014年）129-131頁参照。

あり方、社会のあり方を根本的に変更すること――いわゆる予防国家に向けて大きく舵を切ること――にもなる。さらに、データないし統計によって個人の行動を予測し、例えば潜在的犯罪者としてラベリングすることは、個人を理性的あるいは自律的な行為主体として考えてきた近代国家・近代市民社会のあり方そのものを変更することにもなる。＜適用＞が、個人の尊重原理という憲法の基底的な価値と抵触しうるという先述の指摘は、この点に深くかかわるものである。

このように、警察によるデータマイニング結果の＜適用＞は、個人の権利自由を侵害するだけでなく、国家と個人との関係や社会のあり方そのものを変更する力を有している。そうすると、その正当化には、「我々」自身の同意、別言すれば、安全のために、警察による事前的・予防的介入をも受け容れるという「我々」自身の覚悟表明、すなわち「法律」が必要となるように思われるのである。かくして、集合的同意としての法律は、現実に取得が困難である個別的同意の代替手段としての消極的な意味をもつだけでなく、警察でも裁判所でもなく、「我々」自身が、国家ないし社会の基本的なあり方を決定するという積極的な意味をもちうるのである。

以上のようにみると、警察によるデータマイニング結果の＜適用＞を憲法上正当化するには――侵害留保論的にみても、本質性理論的にみても[69]――まずは「法律」という「形式」が必要になるように思われる。しかし、それに限られるわけではない。第1に、＜適用＞を実質的に正当化するだけの「理由」が求められる。例えば、「万引き」を迅速に捜査するとか、予防するといった「理由」で、＜適用＞は実質的に正当化されるであろうか。ここでその詳細を論ずることはできないが、少なくとも、先述した＜適用＞の侵害的性格や、個人の尊重原理との関係を踏まえて、いかなる場合に＜適用＞を実行することが許されるのかを慎重に検討し、これを法律に明記すべきといえよう。

第2に、情報主体本人によるコントロールに代わって、適切な運用（上記

69）前掲注68）参照。法律の留保論につき、近年の議論をまとめたとして、原田大樹「法律による行政の原理」法学教室373号（2011年）4頁以下参照。

「理由」ないし目的に従った＜適用＞）を担保するための仕組み・構造を組み込むことが求められる。周知のように、いわゆる住基ネット合憲判決[70]は、住基ネットを合憲と判断するための構造的な条件として、①セキュリティ・システムの堅牢性、②懲戒処分または刑罰による目的外利用・漏洩等の厳格な禁止、③監視機関等、「適切な取扱いを担保するための制度的措置」の採用を求め、これらの諸条件が満たされ、住基ネットで取扱う本人確認情報が濫用・漏洩等される「具体的な危険」がないことが、その合憲性の根拠とされた。警察によるデータマイニングやその＜適用＞についても、こうした構造的条件が必要になると思われるが、取扱う情報の性質の違いから、住基ネット判決で示されたものとは異なる、あるいはそれよりも厳格な条件が求められると考えるべきであろう。

例えば、データマイニング結果の個人への＜適用＞は、その者が重罪等を犯すリスクや可能性に関する情報を生み出す。この情報が外部に漏洩された場合の影響は計り知れず、それだけ、厳罰をもってこうした行為を禁止する必要性が高いといえよう。また、＜適用＞の目的外の実行や、＜適用＞結果に対する過剰反応を抑止するために、＜適用＞を行う者、その結果にアクセスできる者を絞り込むとともに、＜適用＞結果はあくまでも統計的な推測に過ぎず、当然に誤差を含むといったリテラシー教育を、関係者に対し徹底して行う必要もあるように思われる。さらに、監視機関の専門性を高めることも重要である（この監視機関をいかに組織するか、国会との関係、国家公安委員会との関係、今後導入が予定されているプライバシー・コミッショナーとの関係をいかに整理するかは、今後の課題としたい）。というのも、データマイニングが公正かつ適切に行われているか、つまり、データマイニングによって抽出・発見されたパターン等が信頼に値するものかなどを判断するには、高度に専門的な知識が必要となるからである。技術者・専門家による恣意は、技術者・専門家でないと見抜けない、というわけである[71]。この点に関連して、マイヤー＝ショーンベルガーとク

70）最判平成 20 年 3 月 6 日民集 62 巻 3 号 665 頁。

キエが以下のように述べていることが注目される。

> 「今のコンピュータが何らかの判断を下す場合、プログラムに記述されたルールに従って処理する。だから、もしコンピュータにおかしな動きが見られたら、プログラムにどういうルールが書かれていたのかをチェックすればいい」。「ところが、ビッグデータ分析の場合、このようにさかのぼって調べることはきわめて難しい。アルゴリズムによる予測は、あまりに複雑すぎて、ほとんどの人には理解できないことが多いからだ」。したがって、「ビッグデータ予測とその背後にあるアルゴリズムやデータセットは、ブラックボックス化する危険性がある。責任の所在も不明だし、さかのぼって調べることもできないから、信頼もできない。そうならないためには、ビッグデータの監視と透明化が必要であり、それを支える新たな専門知識や制度も欠かせない」[72]。

マイヤー＝ショーンベルガーとクキエは、このように述べたうえで、アルゴリズムの専門家――「アルゴリズミスト」――を監視機関内に配置することを提案している。監視機関において、「コンピュータサイエンスや数学、統計学の分野の専門家であり、ビッグデータによる分析・予測の評価役を担う」アルゴリズミストは、「公平と機密保持を旨とし、情報源の選択、分析・予測ツール（アルゴリズムやモデルを含む）の選定、分析結果の解釈について評価」し、問題が生じた場合、「使用されたアルゴリズムや統計手法、データセットを調査する」というわけである[73]。警察によるデータマイニング結果の＜適用＞を認める場合にも、このような専門家を監視機関内に配置し、監視の実質化を図るべきであろう。逆にいえば、こうした専門家不在の監視機関では、先述した、合憲といえるための構造的条件をクリアしない可能性がある。

---

71）技術者、あるいは「コード書き（code writers）」の恣意をどのように防ぎ、彼らの「権力」化をどう統制するかについては、Danielle Keats Citron, *Technological Due Process*, 85 Wash. L. Rev. 1249, 1254-1255（2008）.

72）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）265-267 頁。

73）ビクター・マイヤー＝ショーンベルガー＆ケネス・クキエ・前掲注 31）267-268 頁。

いま述べた監視機関の充実は、ザルスキー（Tal Z. Zarsky）の指摘する透明性（transparency）の限界という観点からも重要である。ザルスキーによれば、データマイニングやその＜適用＞が公正かつ適切に行われるためには、そのプロセスに関する情報を一定程度公開し、透明性を確保することが重要であるが、他方で、警察内部のデータ処理プロセスについては、その全てを公開すればよいというわけではないとされる。例えば、テロを起こすかどうかを予測するパターンや、その各要素（ａ＋ｂ＋ｃ＋ｄ→テロを起こす可能性が高い）を公開した場合、潜在的テロリストは、要素となる行動（ａ、ｂ）を避けたり、要素となる属性（ｃ、ｄ）を変更するなどして、標的化を回避する可能性がある[74)]。また、重罪を犯すかどうかを予測するパターンの要素として、「人種」等が含まれていた場合、その公開は、人種差別等を助長する可能性もある[75)]。このように考えると、警察によるデータマイニングの全プロセスを透明化することは必ずしも望ましくなく、そのうちのいくつかは意図的に曖昧化されている必要がある。言いかえれば、＜適用＞を行う目的・理由や、結果情報にアクセスできる者の範囲、濫用・漏洩等した場合の制裁などは法律にしっかりと明記し、公に開示しておく必要があるが、予測アルゴリズムの詳細等については、むしろ公に対して（は）秘匿しておくべきなのである。そうすると、予測アリゴリズムの適切さを情報主体や公衆に代わって審査する組織として、アルゴリズミスト等を擁する監視機関の役割が非常に重要になるように思われる（その前提として、監視機関に対してはこのような情報が開示されている必要がある。このような手続も法律上明記しておくべきであろう）。言いかえれば、透明性に限界があるからこそ、監視機関の役割が非常に重要になるということである[76)]。法律上、こうした役割に見合った権限が監視機関に与えられていなければ、先述した合憲の条件をクリアしていないと考えるべきであろう。

74）*See* Zarsky, *supra* note 16, at 1553-1560.

75）*Id*. at 1560-1563.

76）*Id*. at 1558, 1562, 1563-1564, 1566.

## Ⅳ　結語に代えて

以上、本稿は、ビッグデータやデータマイニングを用いた予測的ポリシングの手法が、近い将来我が国の警察にも導入されることを「予測」して、その（憲）法的課題がどこにあるのか、なぜそれが問題なのかを考察し、かかる手法に対する統制のあり方を――予備的なものとはいえ――示したつもりである。それによれば、データマイニング結果の個人への＜適用＞（とくに個人の行動予測を目的とするもの）は、単なる個人情報保護の問題を超えて、思想・良心の自由や、自律的存在としての個人を尊重する憲法の基底的原理（個人の尊重原理）にも関連する問題を含むことが明らかになった[77)]。したがって、ビッグデータ時代における自己情報「コントロール」の実際上の不可能性・困難性によって、憲法的規律の必要性をうやむやにすべきではなく、この「時代」に見合った、実効的な統制手法を構築すべきなのである。我が国の警察における予測的ポリシングの動向を注意深く見守りつつ、議論を深化させて行く必要があるであろう。

77）民間企業によるビッグデータ利用やデータマイニングについても、「個人情報保護」の問題が論じられることがあるが、問題の本質はそこにあるわけではない。ビッグデータによる行動予測等が個人情報保護の議論に矮小化された段階で、社会科学的な知は必ず敗北する。マイヤー＝ショーンベルガーとクキエが、「これから我々個人にとって怖いのは『プライバシー』よりも『確率』となる」（前掲注 56）参照）とか、ビッグデータのマイナス面は「我々を確率という名の牢獄」に放り込むことである（前掲注 55）参照）と述べているように、そこで本来論じなければならないのは、「個人の尊厳」そのもののゆくえである。