

## 効率性や経済合理性を「ほどほどに」統御するための公法学——山本龍彦（編著）『A Iと憲法』に寄せて

横 田 明 美

### 1. レビューの対象

2016年から2018年にかけて、法学分野においてもA Iの利活用とそれに伴う法的課題についての本が次々と出版された。その中でも、異色の色彩を放つのが、山本龍彦（編）『A Iと憲法』（日本経済新聞社、2018年）である（以下、本書と略記し、頁数のみで引用する）。本書は「一般書」の皮を被った「専門書」であり、また、執筆陣が若手・中堅の研究者中心であることにも特徴がある。

なぜ評者が「異色」とみるのかは、この本はA Iによって変わる社会を描くだけでなく、一般市民において流布している憲法のイメージをも再定位しようと試みている点にある。本書の「はじめに」〔山本龍彦〕にもある通り、現在の日本では、憲法といえば「9条論」、「改憲論・護憲論」というイメージがついて回り、「過度に政治化」（9頁）された文脈で扱われる。また、「憲法とは国家権力を縛る法規範である」「国家権力が憲法に服する」という一面的な見方が流布しており、それだけではない憲法の役割、つまり「国家権力に対して、憲法原理が経済市場においても守られるように積極的に『動く』ことを義務付けてもいる」という側面が、見過ごされがちになっている（8頁）。本書はこのような状況に警告を与えつつ、効率性や経済合理性を「ほどほどに」統御するための憲法論を説き、憲法原理と調和したA Iの利活用の方向性を考える（9頁）、というのである。

「上手く実装化すれば憲法原理のより良い実現に資する」A I技術について、過剰に利活用の側面ばかりが強調される議論にも、また、過度に脅威論を煽る議論とも一定程度の距離を取ろうとしている。また、上述の通り、国家による利活用だけでなく、私人による利活用にも言及して、「社会的排除や民主主義の破壊を許すようなディストピア」が、「効

率性や経済合理性」の追及の果てにやってくることを防止するための憲法学を描き出そうとしている。

そこで、本書は序章において、「予測力」という観点から、AIシステムがこれまでの高度情報システムと何が違うのかについて説明した上で、第1章では憲法原理としての個人の尊重、プライバシーとの関係（ここまで山本龍彦）を検討し、それ以降の章では、各領域における「利用」が、どのような影響をもたらさうかを論じている。各章の冒頭には、想定される利用のあり方とそのリスクについて、読者とイメージを共有しながら議論を進めていくために、「リスクシナリオ」（仮想の事例問題）が用意され、理解を助けている<sup>(1)</sup>。

行政法学を出発点として環境法・情報法を研究する評者からすれば、本書が示そうとする枠組みは、まさに「AI時代における公法学のあり方」であるといえる。行政法は、本書との関係では「憲法実現法律」・「憲法具体化法律」（8頁）と捉えられる個別法を取り扱う隣接領域であり、公の秩序を実現するための法として、憲法学と同じく公法学に属する。評者はかつて、行政規制と行政による利活用についての考察を行ったことがある<sup>(2)</sup>が、本書が示すのは、より理念的な、それゆえに見過ごされがちな「マクロ」の制度設計上の課題である（なお、対応するはずの「ミクロ」の個別救済は、非常に難しいことが本書においても明らかになる）。

これまで、「地域の法と政治研究会」においては、政治学・行政学・行政法学の研究者が、地域の自治に影響を与える理論について議論を進めてきた。その観点から、この書評において、AIシステムが利活用された世界における公共性の観念を中心に上げることとする。AIが利活用されていく社会において、とりわけ政治との関係で、何が起きうるのかを法学の観点から論じた書籍は、まだ多くない。「はじめに」の背景にあるとおり、本書は全体としてビジネス面への利活用を念頭においた記述が多いが、だからといって、公的部門や、政治過程における利活用について論じていないわけではない。そこで、以下

---

(1) このようなリスクシナリオを利用した議論はしばしばなされており、福田雅樹・林秀弥・成原慧（編）『AIがつなげる社会』（弘文堂、2016年）でも用いられた。なお、本書第1章のリスクシナリオは、同書に収められた山本龍彦「AIと『個人の尊重』」でもちいられたものと同一である（本書62頁）ことから、本書は福田ほか（編）の手法を受け継ぎつつ、憲法学の観点からより深掘りしたものと位置づけられよう。

(2) 行政規制について、横田明美「ロボット・AIと行政規制」弥永真生・宍戸常寿『ロボット・AIと法』（有斐閣、2018年）103-130頁、行政自身によるAIの利活用について、横田明美「行政によるAIの利活用と行政法学の課題」自治実務セミナー2019年1月号9-13頁。

では、本書の全体を通覧する序章：A Iと憲法問題、第1章：A Iと個人の尊重、プライバシー〔山本龍彦〕と、政治過程に関連の深い章として、第6章：A Iと民主主義〔水谷瑛嗣郎〕、第7章：A Iと選挙制度〔工藤郁子〕について、本書が提起した問題点の概観と本書が提示する解決策を中心に紹介し、それに基づいて行った研究会の議論を踏まえて、若干のコメントを付することとしたい<sup>(3)</sup>。

## 2. 本書が提起する公法学的諸問題

以下では、序章、第1章、第6章、第7章が提示する、公法学の見地からみたA I利用の進展によって生じうる諸問題について紹介する。

### (1) A Iシステムの利活用と「予測力」（序章・山本）

序章において示されるのは、A Iの「予測力」の内実と、その権力性である。

「プロファイリング」（19頁）とは、A Iは過去の膨大なデータから、人間には予想も出来なかった事物と事物の相関関係やパターンを発見・認識し、この関係・パターンを特定の個人に関するデータセットに適用することで、当該個人の趣味嗜好、健康状態、精神状態、政治的信条、職務上の適格性、信用力、学力、知能、収入などを自動的に分析・予測できる事を指す。メリットとしては、個別的（personalized）な情報配信が可能となることであり、これは広告のあり方を大きく変えるものである（参照、第2章：人間の自己決定への影響）。また、ニュース配信や政治的利用も想

---

### (3) 本稿で紹介を省略した各章・各コラムのタイトルと執筆者は次の通りである（登場順）。

コラム：A Iとデータポータビリティの権利〔生貝直人〕、第2章：A Iと自己決定原理〔古谷貴之〕、コラム：A I利用の「風潮」と自己決定〔瑞慶山広大〕、第3章：A Iと経済秩序〔市川芳治〕、第4章：A Iと人格〔栗田昌裕〕、コラム：A Iと尊厳〔河嶋春菜〕、第5章：A Iと教育制度〔堀口悟郎〕、コラム：A Iと知的財産〔太郎田耀〕、第8章：A Iと裁判〔柳瀬昇〕、第9章：A Iと刑事法〔笹倉宏紀〕、解説：我々とは誰のことか、あるいはA Iという他者〔成原慧〕。執筆者が比較的若い（40代前半から20代後半）ことだけでなく、主領域が憲法学あるいは情報法学である者と、利用が想定される分野を主領域とする者とが混在している点も興味深い。特に後者については、しばしば当該分野では見過ごされがちな、「憲法学」と当該分野との関係が明示的に語られる記述がある（例として、第3章：A Iと経済秩序〔市川芳治〕では、憲法それ自身が「経済秩序」について何も語っていないことから、各国の競争法が「経済憲法」などと呼ばれていることを紹介する（166頁））。

定される（参照、第6章：民主主義）。

そして山本は、Step 1：大量のデータが収集され、プールされる「ビッグデータの収集・集積」、Step 2：人間では気づかない相関関係や行動パターンの抽出による「解析」、Step 3：相関関係やパターンが特定のデータベースに適用され自動的に予測される「プロファイリング」、Step 4：予測結果が特定の目的のために「利用」され、Step 5：予測結果の妥当性検証のための事後的な「追跡」という5つの段階が循環するデータの循環（データサイクル）が一般化すると指摘する（20-21頁）。

このような予測は、既に米国のシカゴ警察における予測的警察活動（predictive policing）として実用化されているところ、ここで「予測力」の権力的側面が問題となる。確率的予測は本人そのものを見たわけではない。あくまで「共通の属性を持った集団」（セグメント）の一般的傾向をみてフラグを立てているのであって、AIの予測は「確率」を示すにすぎない。しかしこの確率的予測は、まったく重罪犯罪を行う気のない者に対しても「力（power）」としてのしかかる。もっとも、これはAIを利用する側にとっては極めて「合理的」であるところ、ここでの「合理的」とは、経済的合理性か、それとも憲法的合理性かが問題提起される。また、ここで「効率性」と「個人の尊重」とのバランスの取り方が、憲法的問いとして析出される（23-24頁）。

また、もうひとつの権力性の源泉として、「理由」が説明されないということが指摘される。理由がわからなければ、その「評価」に対して有効な反論を行うことも難しいからである。そうして、否定的評価が人生にずっとこびりついていくことになる（24-25頁）。これは、とりわけ個人を対象としたAI予測を行政過程において利用する際に放置できない課題である。既に行政手続においては行政手続法において理由の提示が申請に対する処分（8条1項）・不利益処分（14条1項）のいずれにおいても必要とされている。それは、行政手続における公正性・透明性の向上をはかり、不服申立てに便宜を与える趣旨であるところ、アルゴリズムの複雑性や理解困難性はそれらを阻害する可能性があるからである（36-41頁）。

AI利活用が政策立案や立法過程において進んでいく際に、「代表性」との関係が問題となる。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ゆがんでしまえば、容易に差別や誤りが入り込む要因となる。ここでは、過少代表（underrepresentation）問題が米国FTCの報告書における事例を元に紹介される。補修が必要な箇所を発見するための道路状況調査に

においてスマートフォンからの位置情報データを元にしたところ、高所得者層居住エリアの住民達が過剰に代表され、低所得者層の居住エリアデータがデータ・プールに過少に代表されてしまったという事例である（34-35頁）。また、データを取られたくない人たちから「集めてはならない」要請との対立も懸念される。「データ・プールに『現れない』者はこの世の中に『存在しない』者として認識する」ため、データ・プールの「偏り」は、少数者への差別を再生産ないし助長する可能性がある（44頁）。また、歴史的差別やバイアスもデータ・プールの偏りをもたらすため、現に存在している差別を新たな決定に組み込むリスクを内包することになる（45頁）。

## （2） 公法原理との対立 — 「個人の尊重」原理、プライバシーとの緊張関係（第1章・山本）

第1章では、序章で説明されたAIの「予測力」が活用される就職活動を想定したリスクシナリオのもと、「個人の尊重」原理とプライバシーの権利との関係が議論される。

まず、憲法上の個人の尊重原理が、①人間の尊厳→②狭義の個人の尊重（集团的拘束からの自由）→③個人の尊厳（自律）→④多様性・個別性の尊重という4層から成り立っていると分析される。そして、前提条件である①と尊重された結果である④の中間にある、②と③が個人の尊重原理の核心であるとする。つまり、②「個人が人格的存在として平等に尊重されること」、③「個人が自律の能力を持つことを前提に、誰からも命じられることなく、主体的に自己の人生をデザインしていくことを認めさせる、という積極的側面」が、個人の尊重原理の核心であるというのである（64-66頁）。

これに対して、AIの予測、特に上述のデータの循環（データサイクル）におけるStep 3では、具体的な「個人」そのものを扱っているのではなく、「共通の属性を持った集団」として、つまりセグメントとしての一般性を取り扱っていることに注意が向けられる。様々な背景や経路を持った複雑な属性を「丸める」ことで、コンピュータにインプット可能な形に抽象化されるからである。そこでの予測は、「具体的存在としての個人」は、どこまで行っても「属性の集合（セグメント）には還元されない」ことから、個人として評価すべきとする個人の尊重原理と根本的に矛盾する可能性がある（67-68頁）。もっとも、後述するように、プライバシーや自己情報コントロール権との関係ではStep 1が「特定の個人には関心が無い」ことが重要となる

(88頁)。

A Iによる予測と本当の自分とが乖離していることに対して反論し、結果を訂正する難しさ(69-70頁)は、コンピュータによる自動化された判断を過信してしまう認知的傾向にあらがえないという「自動化バイアス」、A Iを利用する側も「何だかよくわからないけれど」と言わざるを得なくなる意思決定プロセスのブラックボックス化によって説明される。いったんA Iに「嫌われ」た者が挽回の機会を与えられることなく「理由」も知らされないままに排除され続けるという「バーチャル・スラム」(71-72頁)への有効な反論手段はない。従来の人間による判断は利用する「情報」が限定されていたため、本人も理由が推測しやすいが、同じブラックボックスでも、多様なデータを基礎にしたA Iの意思決定のほうが「はるかにその闇は深い」と評価される(73頁)。

それでは、予測と現実との乖離を少なくするにはどうしたらよいか。データ量と予測精度は比例関係にあるため、「もっとデータを」(“More Data”原則)という考え方がA I社会において一つの基本原則になる(74頁)。しかし、この“More Data”原則もまた、個人の尊重原理との緊張関係が生まれる。まず、人生をやり直す自由(更正する権利)については、個人の尊重原理に含まれると論じる本書の立場からは「忘れられる権利」あるいは消去権を、一定の条件でデータ・スティグマを洗い流す権利として位置づけ直す必要がある、と説かれる(77頁)。次に、データを利活用することによって予測の精度を高める必要性と、憲法がこれまで正面から否定してきた生まれつきの特性(遺伝情報)など、本人自らでは修正できない要因による差別を復活させてはならないという議論との衝突が指摘される。これを山本は、データサイエンス・保険数理学における「公正」と、憲法の観点から見る「公正」の衝突と表現する(78頁)。

自己情報コントロール権として理解されるようになったプライバシー(84-86頁)を前提にすれば、「A Iに情報を与えたくない」と考える者が行使する自己情報コントロール権は、A Iの成長を妨げるという意味でも、またデータ・プールにおける偏り(上述の「過少代表」問題)を引き起こすという意味においても、より多くのデータ(more data)を必要とするA I社会と矛盾する(A I社会における「プライバシー・ジレンマ」)。ただ、データ収集段階(Step 1)においてはセグメントとして、個人が特定されない形でのデータを「集合界」にプールすれば、プライバシー権ないし自己情報コントロール権を直接侵害することなくA Iにより多くのデータを供給で

きる。しかし、より多くの属性情報を集積すればするほど、それだけ個人特定のリスクが高まる（88頁）。

また、個人への適用（「個人界」）段階であるStep 3において、プロファイリングを通じた精度の高い「予測」は個人情報の取得を迂回する状況（迂回的取得）をもたらす。これは、これまで個人情報保護法が要配慮個人情報に本人同意を必要としていること（同法17条2項）との関係が問題となる（90-91頁）。これまで人の頭の中で行われてきた予測・推知について本人同意が問題とならず、なぜAIを用いた予測（プロファイリング）がプライバシー上問題となるのかを議論しなければならないからである。この点で、山本は『宴のあと』事件判決（東京地判昭和39年9月28日下民集15巻9号2317頁）が「私生活上の事実らしく受け取られるおそれのある」情報をも保護されるべきとしたことを参考にして、精度の高い要配慮個人情報についての予測のうち、コンピュータ上に出力されたとき、社会通念上、人に「事実らしく受け取られるおそれ」のあるものについては、要配慮個人情報の「取得」（17条2項）として扱うべきであり、それゆえ事前の本人同意が必要となると主張する（92-93頁）。

### （3） AIと民主主義（第6章・水谷）

第6章冒頭のリスクシナリオは、AIによる「代理人」・AIによる政策立案と、AIによる情報環境構築に関わるものである。

もしAIが「人（政治家）」よりもはるかに効率良く、裏切らず、主権者たる国民の「代理人」を務めるとするならば、どうなるだろうか。第6章冒頭で示されるリスクシナリオ〈未来予想図1〉は、「人民の意見を政治の場に集積することを自動化する意見集約型のAI」が、AI議員禁止法案について人間の議員と議論するものである（286頁）。実際にプロジェクトが進められている「ロボット大統領」構想（294頁）を前提にそれを導入するならば、国民主権原理を維持するためには憲法改正による導入と更なる憲法改正による代議制復活可能性が残されていることが必要となる。

代議制の廃止というラディカルな移行をもたらすのではなく、よりソフトな方法、つまり「大衆の無意識の可視化」を行い、代議制を補うものとして用いてAIにサポートされる政治を目指すという方向性も検討される。実際に、この章が執筆された2018年4月当時、多摩市長選の「松田みちひと」候補が、市議会の議事録をAIに読み込ませ、多摩市における社会問題を抽出するなどのAI政治活用を進めると主張して選挙戦を戦ったこと（ただし落選した）が紹介されている（318頁注3および319頁

注5)。

また、データ駆動型の政策判断を行えるファシリテーター（政策提言者）として、政策立案のためにAIを用いる方向性は、総務省AIネットワーク化検討会議の報告書や、NHKの「AIに聞いてみた どうするのよ!? ニッポン」という番組においても示されたところである。実際、NHKの番組で用いられたAIが、「40代ひとり暮らし」が増えると「自殺者数」「餓死者数」「空き家数」「救急出動回数」等が増え、「合計特殊出生率」「老人クラブ会員数」等が減るという傾向を示したことが話題になり、荒川区が対策の必要性につき検討を始めたことが紹介される（296-297頁）。

このようなオープンデータ分析結果の活用と政策立案への利用に対して、水谷は「制度化された熟議」の後退が生じ、政治家の役割が後退するという懸念を示す。フォーマルな決定志向の審議過程とインフォーマルな意見形成（熟議）における議題提供の両者の協働による民主的「正統」性が確保されるという理解を前提にした「制度化された熟議」（298-299頁）と、AI利活用の関係が問題となるのである。

NHKの番組で使われた「AI」は、導出したパターンがなぜそうなるのかの「理由」が示されない、あくまで確率的・統計的に正当化された相関関係のみを示すのみであるため、人間による解釈としての因果推論が必要となる。「ラブホテル軒数」が増えると「女性のボーナス額」が増えるというような「全く因果関係の読み解けない」パターンすらある（300-301頁）。

さらに続いて、AIのブラックボックス化がもたらす意思形成過程のブラックボックス化について、直接的な国家意思形成過程に大きな疑念が投げかけられることを示している。「AIの選択が何をもってなされたか、人間の誰にもわからない」、「AIに振らせるサイコロの目を信じて賭けた、本質的にはそれだけだ」、「昔から占いや霊能力を信じた政治家は数多くいるが、その新しいバージョンでしかない」——水谷が印象的に用いるこれらの印象的な表現は、人工知能学会の研究者と小説家とがコラボレーションすることで生まれたアンソロジーのうち、政治過程を扱った長谷敏司の小説<sup>(4)</sup>において、決断をアウトソースすることにした政治家の独白の引用である。

---

(4) 長谷敏司「仕事がいつまでも終わらない件」人工知能学会（編）『AIと人類は共存できるか？ — 人工知能SFアンソロジー』（早川書房、2016年）327頁。不都合を握りつぶす過程、24時間計算し続ける機械に合わせるために使い潰される人力、そして数値で示される国民投票に振り回される政治過程……同小説は非常に多くの示唆を投げかけた、それ自体が一級の「リスクシナリオ」となっている。

そして、水谷はこのAI利用がもたらす変化を、かつて諮問的レファレンダム（「数としてしか示されない民意」）に対して示された懸念を参照して分析する。政策立案にAIを「安易に」用いれば、政治家同士あるいは官僚や選挙民との間における「理由を探るコミュニケーション」が後退し、他者の観点を考慮に入れることでもたらされるはずの「選好の変容」も減少していくこととなる。さらには、政治家の説明責任の関連からは、失敗した際の自己反省とトライアンドエラーが不可欠であり、AIの計算間違いやデータセットの不足以上の説得的な説明が求められることを指摘する（301-303頁）。

もう一つのリスクシナリオ〈未来予想図2〉は、AIによる「オススメ」によってインターネット空間を流通する情報環境の「個人化」（personalization）が進む事に対する懸念を示している。これが、「公共圏」の崩壊をもたらし、「集団極化現象」を加速化させる恐れがあるからである。人々が自分好みの新聞を選ぶことで生じるとしてかつてから指摘のあった「私のため」に「私が選択した」情報環境（「デイリーミー」）と、パーソナライズ・フィルターに囲まれている現在生じつつある「あなたのために」に「AIが選別をした」情報環境（「デイリーユー」）の違いが民主主義に与える影響は深刻である（304-309頁）。新聞であれば、同時に自分と同じ価値観を持つ人が読むという意味で「一般化された新聞」であるところ、自らが関心のある（と推測される）個々の情報にだけ触れ、そうでは無い情報には触れることがないという個人化された情報に接することによる孤立化が進展する。しかも、このパーソナライズ・フィルターは情報の受け手に見えにくく、「偏り」を認識できないために、「偏向のない客観的眞実」だと信じ込まされてしまう。そして、編集者の意図が介在せず、AIが「あなた」のために勝手に推測するために、人々は情報をただ受け取るだけの受動的立場への変更を余儀なくされるからである。これらは、民主主義が前提にする「私たち」という観念そのものを分断し、他者との討論を通じた真偽判断ができなくなり、「思想の自由市場」が機能不全になること、受動的立場を利用した政治動員、誘導の危険……というかたちで、能動的な機会の喪失と制度的意思決定（権力行使）に対する民主的「正統性」の掘り崩しが生じる。

本章で示された「民主主義における意思集約過程あるいは政策形成過程の自動化」と、既に生じつつある「情報生成・流通の自動化」（317頁）が、ロボットによるニュースの自動生成等によりさらに加速する（311-316頁）未来が予言されている。

#### (4) AIと選挙制度（第7章・工藤）

第7章では、政治的に用いられる「フェイク」合戦（326頁）がリスクシナリオとして提示される。これは、2016年米国大統領選挙で話題となったケンブリッジ・アナリティカ事件をモチーフにしたものであり、それは端的にいえばパーソナル・データを利用した政治マーケティングを巡るパーソナル・データ不正取得と世論誘導疑惑である（330頁）。本章では、ハンス・ケルゼンの価値相対主義を前提にした民主政理解が、今後修正や変容を迫られるのかが問われている（328－329頁）。

米国における実例として紹介されているのは、マイクロ・ターゲティングと得票予測、そして「推知」による個人情報取得の潜脱（331頁）である。これらは第1章で詳述されているところの具体例となっている。そして、第6章の具体例として、政治的メッセージのパーソナライズ（個人化）（332頁）と、マス・マーケティングが後退することによる公共圏の崩壊が示される。

このように、本章では、第1章・第6章の指摘に具体例を与える記述が多い。フェイクニュースの実態を紹介する箇所では、米国のトランプ陣営有利のフェイクニュースの出所が遠く離れたマケドニアだったこと、そしてフェイクニュースのエンターテインメント産業としての側面（336－337頁）が紹介される。それらを根拠として、「思想の自由市場」論が想定するような自由競争による虚偽・低湿な議論の淘汰が働かないことによる機能不全と、AIによってフェイクニュースが量産されることによりファクトチェックが今以上に追いつかなくなっていく懸念が示される（337頁）。

ここで、「思想の自由市場」論が想定している「自由と幸福の一致」と「強い個人」という二つのフィクションが確認される。すなわち各人が評価能力や判断力を備え、一定の自律的な判断を行いうるという人間が、各人の幸福が何であるかを知っているからこそ、自由を確保すれば自ずから最大の幸福が実現される——この個人の自己決定と集団的決定の自律性はある種の擬制であり、自然的事実とは異なるからこそ、規範的に制度を構築することが必要となると工藤は整理する（338頁、この箇所については後掲4.(2)で引用する）。

### 3. 本書が示す対応策

それでは、これまで示された課題に対して、何らかの打つ手があるのだろうか。以下で

は、本書自体が示している対応策を紹介する。

### (1) 自己情報コントロールの支援とGDPRにおける「権利」(第1章・山本)

“More data”原則とプライバシーの緊張関係を示した第1章では、その対応策として、自己情報コントロールの支援を目指した制度設計が紹介される(94-98頁)。市場支配的地位の濫用としてプラットフォーム事業者による個人情報収集を独禁法の観点から規制すべきか否かという検討、個人情報の開示先を本人の選好に従ってコントロールしてくれる情報銀行制度、別の事業者へ情報を預け直す権利としてのデータ・ポータビリティ(データ可搬性)の議論が、個人の主体性を回復させるための重要な制度として紹介される。

また、データ・ポータビリティの議論は、2018年にEUにおいて施行された一般データ保護規則(GDPR)から火が付いた議論でもある。GDPRがデータ主体の権利として認めている異議を唱える権利(中止請求権)(21条1項)と、自動処理のみに基づき重要な決定を下されない権利(22条1項)とそれに関連するガイドライン(100-102頁)を紹介した上で、山本は、AIと人間とが両者の短所を補いながら協働する「ケンタウルス」モデルを基本的人権の保障という観点からGDPRが要請しているとして、これらの規定が個人の尊重原理に関する上述の分析と親和的であるとして、高く評価する(105頁)。

### (2) プラットフォーム事業者の役割と「メタ自己決定論」(第7章・工藤)

フェイクニュースを巡る諸問題については、プラットフォーム事業者に、ゲートキーパーとしての役割を期待する動きがある。情報媒介者であり広告配信主体でもあるIT事業者に対して、法規制あるいは自主規制による対応を求める動きが進んでいる。プラットフォーム事業者へのフェイクニュース規制(339-340頁)として紹介されるこれらの動きとともに、様々な試行錯誤の事例が紹介される。例えば、フェイクニュースの登場順が下になるようにするアルゴリズムに対する技術的介入(341頁)だけでなく、表示回数コントロールによる投票促進など、政治への関心を高める啓発も行われる。

ここで問題になるのは、これが好ましい「ナッジ」なのか、それとも忌むべき世論操作なのか(342-344頁)である。「ナッジ」とは、個人の選択の自由を尊重しつつ、個人が「より良い」選択を行えるように後見的に介入することである。他方、これら

の実験や機能の導入を、恣意的な選挙区割と同様の「デジタル・ゲリマンダー」として批判する見解もある。つまり、情報環境の操作は、恣意的な選挙区割り操作と類比可能だということである。工藤は、この対立を紹介しつつ、選挙制度に関する日本国憲法の規律密度が低いことを指摘する（344頁）。憲法前文の示す選挙の「正当」性及び43条1項・2項が示す選挙の方法をみても、前者が望ましい選挙像を具体化していないため、後者の限界を画することができないという。

また、これらのアーキテクチャを介した働きかけと自己決定権（憲法13条）の関係も指摘される。知らないうちに自由・自律が浸食される事態を避けるために、アーキテクチャ自体の選択可能性それ自体を確保する「メタ自己決定論」によるアプローチが紹介される。もっとも、それをいかなる形で実現するのか、誰が行うのか、不当な場合の是正手段や強制執行の可能性など、実装の困難性も合わせて指摘されている（344—345頁）。

## 4. 若干のコメント

以下では、AIに関する法制度設計について研究する者として、本書の通奏低音というべき考え方について、研究会における討議も紹介しながら、若干のコメントを付すことにする。

### （1） 困難に晒されている公法原理

#### ① 情報環境の変化と個人・集団（セグメント）

本書評で取り上げた範囲に共通して見られる危機感は、「個人」と「集団」の緊張関係である。個人の尊重原理と集団として扱われることが、どのような関係にあるのかということ鮮やかに映し出している。

第1章において個人の尊重原理との関係で議論されていることの核は、データの集積・分析において人が集団（セグメント）として扱われることそれ自体である。個人に紐付けられないかたちで、個人の同意を前提とせず集められた情報が、集団を経由したうえで個人に「予測」として当てはめられるという一連の過程を経ることで、結局は個人のそれぞれの生活に影響を与えていく。そうすると、リアルな個人がデジタル世界に正しく写し取られることだけでなく、「適切に予測される」

環境を構築すること自体も求められることになる。しかし、公衆衛生の確保のためには個々人の「予防接種を受けたくない」という希望が障害になりうるのと同様、情報を提供したくないと望む個人の希望が、結果としてこの「適切な予測」を掘り崩すことになる。

このような観点は、「デイリーユー」というキーワードで代表されるような、情報環境の個別化・個人化を問題視する第6章・第7章にも共通している。「ユーザーのために」最適化された情報流通環境は、全体としては社会階層の分断と民主主義の危機をもたらしかねない。技術の進展によって可能となった個別最適は、必ずしも全体最適を意味しない。

## ② 「プロセスが明らかにされないことがもたらす変化」——理由と異議申立て

もう一つの通奏低音は、「プロセスが明らかにされないことがもたらす変化」である。これは逆に見れば政治過程・政策過程においてプロセスがなぜ重視されるのかということの意味を問い直している。「ヒトが理解可能な形で説明できるのか」という問いは、ヒトとAIシステムの関係を示しているが、理由の提示が必要となる場面で重要視される「ヒトが異議を唱えることが可能であるために」というときの「ヒト」は、上述の「ヒト」と同じ者とは限らない。例えば行政がそのようなシステムを利用したというときに実際に適用されるのは間接利用者にあたる「ヒト」である申請者であり、AIを使っている「ヒト」である行政担当者（最終利用者）にとっては理解できても、申請者には上手く説明できないという状況も想定できる。つまり、単に機械との関係が第一次的な「ヒト」だけではなく、第二次的な「ヒト」も想定された問いにならざるを得ない。本書は、公法学の観点から二重のブラックボックス状況<sup>(5)</sup>を克明に描写している。

デジタルへの過剰な信頼、何が評価され何が評価されないのかわからないために生じる過剰な内面化、過去の評価がつかまとうことの恐怖——これらは、プロセスが明らかにされないことがもたらす変化として描写される。

もっとも研究会の議論においては、本書が政策決定の過程として描写している議

---

(5) 横田・前掲注(2)「行政によるAIの利活用と行政法学の課題」12頁では、「行政内部の意思決定過程において、AIサービスを提供することを業とするプロバイダーとAIシステムを利用するのは行政の担当者である」ことを踏まえ、行政内部関係では一応説明が仮にいたとしても、行政外部関係として登場する間接利用者である入所希望者との関係では、もう1個の「ブラックボックス」が生じてしまうことを指摘した。

論と、個別決定である行政処分における理由の提示の議論は同一視できるのかという観点から、次のような素朴な疑問が呈された。個別決定の場面においては、行政の恣意性を排除できるという観点（羈束性）を向上させるものとして、むしろ「非人格的な処理」のほうが理想的なのではないか、という疑問である。要件と効果が法定されている行政決定においては、AIを用いた方がかえって恣意性を排除できるのではないか、ということである。しかし、この点については、本書自体が示している懸念として、「数値化する」こと、「集団として取り扱う」ことそれ自体になじまない価値がある場合、「非人格的な処理」に依存しすぎること、それを見落とす可能性があることが指摘できよう。行政裁量における個別事情審査義務が議論となる背景には、数値化できない事情を見落とすことが許容できない場合がありうるからである。これに対しては「個別事情考慮という観点も含めてAIが考慮できるようになればよい」という批判もありうるだろう。しかし、本書はそのような議論に対して先回りするように、“More Data”原則が行き着きかねない陥穽を示している。つまり、データ化されないことを選択した者、あるいはデータ化されにくい状況に追い込まれた者は、その存在や特性を顧慮されなくなるという危険性である。現実世界を「正しく」写し取られないのであれば、予測や適用の場面においてゆがんだ結論が生じ、結果として不利益に取り扱われるおそれがある。

ただ、研究会におけるこの指摘は、EU法制をどのように評価するかという観点からも重要であると考えられる。上述のとおり、EUのGDPRは異議申立権に並んで、「自動処理のみに基づき重要な決定を下されない権利」を法定したうえで、いかにそれを実現するかに腐心している。しかし、それは何のためなのか。判断の責任帰属の問題なのか、それとも「ヒトはヒトによって判断されるべきだから」なのか、それとも「ヒトに理解可能な形であること」を確保しようとするからなのか。データ保護を人権として取り扱うことが謳われているEU法制についての理解を、形式的にだけでなく、実態も踏まえた上で実質論として深めていく必要性を、評者は改めて感じさせられた。「自動処理のみに基づき重要な決定を下されない権利」は、研究会において示された——しばしば行政によるAI利活用に関する市民の反応として示されることもある——「信用できない人間よりも一見信用できそうなAIのほうがマシ」という世論によって否定されかねない理念なのか、それともこの権利が存在することで生じる余波も含めて制度設計上必要なものなのか。今後、AIを巡る法制度・自主規制の枠組みを考えていくうえで、検討すべき課題である。

### ③ 相関関係と因果推論、政策決定における「不都合な真実」

AI が示した相関関係について、人間が解釈により因果推論を行わなければならないこと、それを経なければ因果関係のない議論を根拠にしてしまうことになるため、政策論には用いることが出来ない——このような言明が、どこまで妥当するのかという観点からも、本書及び本書が参照した文献は興味深い。本書第6章で紹介されている2つの事例、(a)「40代ひとり暮らし」が多ければ「出生率」等が下がるという相関と、(b)「ラブホテル軒数」が増えれば「女性のボーナス」が上がるという相関については、それを人間社会の側がどのように受け取るかについて研究会では議論がなされた。

本書では(a)について「『(ある意味で)期待通り』の結果を出した」と評した上で、あえて「この『思いもよらぬ』試算を参考に」荒川区が対策の必要性を検討したことを紹介している(297頁)。この『』の用い方に著者(水谷)の見解が透けて見える通り、この推論は人間においても理解可能である。しかし、これを正面きって政策課題として取り上げることが一種のインモラルに感じられたために、取り上げる事が出来てこなかったのではないか。研究会の議論では、このような「AIに言わせる」ことの意味について議論がなされた。

このAI番組を担当した神原一光ディレクターの言として、「もっと異なる視点が必要なのではないか。もしかしたら、もはや人間だけでは難しいのかもしれない、という思い」が紹介されている。しかし、少なくとも(a)については、人間に難しかったのは、異なる視点そのものではなく、ある種の反発を踏まえてもなおそれを政策課題として取り上げることだったのではないだろうか。つまり、「AIに言わせる」ことで、関係当事者の理解を得ようとする利用法である。これは、コンピュータによる自動化された判断を過信してしまう認知的傾向にあらがえないという「自動化バイアス」(69-70頁)と相まって、次の(b)における破滅的な未来を引き起こしかねない、極めて危険な傾向ではないかと評者には感じられるところである。

ラブホテル軒数と、女性のボーナス。この二つを因果関係として説明するのは極めて困難である。それゆえ、著者(水谷)は「安易に」AIが政策立案に用いられれば、政治家が担うべき「理由を探るコミュニケーション」が後退する懸念を示している。ここで想定されている課題は、既にAIが政策立案に使われることが社会的に一定程度承認されたあと、人間には理解できないレベルでの推論が出てきたと

きに、それをどう扱うのかという問題である。因果推論が成り立たないから採用しないこととするのか、多少強引なこじつけによって「因果推論」が行われたように取り繕うのか、それとも「A Iに言わせる」ことで信用を獲得してしまうのか。結局は、出てきた結果をどのように用いるのかを決めるのは人間であり、その判断の透明性の担保が問題となる。この点、前掲・注(4)で紹介した小説・長谷敏司「仕事がいづまでもたっても終わらない件」では、政治過程と人工知能による予測の相性の悪さが描写される。政権維持を目的として改憲国民投票に対する世論の予測をするために運用されている人工知能が、内閣総理大臣の辞任と国民投票の白紙化を最適な手段として出したときに、ある登場人物が述べるセリフが示唆的である。「不都合な結果は握り潰される。握りつぶせない道具なんて、政治じゃ使えない」と<sup>(6)</sup>。本書第6章では、「A Iの政治利用に関する倫理綱領」を作る必要性が説かれ、その内容として因果推論の検討のための「熟議」期間設定や透明性担保のための監視制度、A I利用の禁止領域（聖域）等が提唱されているところ（303頁）ではあるが、この倫理綱領がどのようなかたちで実効性を持つのかは未知数であると感じられた。

#### ④ 分断による公共圏の崩壊と現実社会

またマクロなプロセスとしては、熟議のプロセスが分断により破壊される懸念が何度も繰り返し主張される。この「異なる意見とふれあうこと」というのがどのような意味を持っているのかということは、失われてみてから初めて分かるということもあり、「民主制の維持」が問われ、見えない形で公共性の前提が掘り崩されていることが指摘されている。分断されていること自体に気づかないままそれが反映されない社会が既に生じつつあることが指摘されている。

もっとも、この点についても、研究会においては疑念が呈された。「熟議」モデルそのものに対する疑念も多く寄せられたが、評者の力量不足であり、それ自体をここで紹介することは差し控えたい。

研究会での率直な議論の中で、最も評者を悩ませた批判は、「ここで掘り崩される公共圏や思想の自由市場論とは、つまりはあるべき理想をうつしたものであって、現実社会とは整合していないのではないかと、少なくとも日本社会をうつしたものととはなっていないのではないかと」という批判であった。確かに、本書のベースとなっ

---

(6) 前掲・注(4)244頁。

ている議論や事例は、その多くが米国を想定した議論であり、米国憲法学の影響を強く受けたものである。また、「日本の法案策定過程を前提にした議論なのかどうか」という問いも、あわせて発せられた。

この点研究会においては、評者は本書も日本の政治過程を前提にしていることを示すために、第6章の「A I の政治利用に関する倫理綱領」（303頁）に付された脚注16（321頁）を紹介した。著者（水谷）は内閣提出法案の策定過程と与党内審査、国会における委員会審議の過程を踏まえた上で、「A I を参照した内閣提出法案に対する対案提出・批判を行うために、与野党の各政党がそれぞれ独自のA I を導入するという構図も想像できなくはない。その場合、『A I 同士の権力分立（内閣A I 対国会・政党A I）』という視点も重要になるだろう」という視点を打ち出している、と。もっとも、いま研究会議事録を読み返してみると、これは米国か日本かという国を超えた適用可能性というよりは、本書全体として「憲法」をどのように語るのか、という観点において議論されるべき疑念であると考えられるので、次項において触れることにしたい。

## （2） 憲法原理のより良い実現とは

今一度、本書全体の企画趣旨である「はじめに」に立ち返り、本書の役割と、目指した価値について触れて本稿を終えることとしたい。一法学者としての視点を離れて、ひとりの現代人として様々な便益を享受している者としては、「個別化」されることの便益や、これまで発見されてこなかった、気づかれなかった問題への対処がなされることも捨てがたい。本書は冒頭（6頁）の「両眼主義」にもあるとおり、サービス自体が「個別化」されることを否定してはいない。むしろ「うまく実装すれば憲法原理のより良い実現に資する」とする。公法学が積み上げてきた「産業の発展と人間性（個人の主体性）のバランス」（4頁）が、空洞化しないようにするための観点を提示したのが本書の役割であろう。

その観点からすれば、今後も「憲法原理」は支持され続けるかが問題となる。しばしば指摘される、「公正」の対立は悩ましいところである。果たして、「憲法学の公正」は、多くの市民に受け入れられるだろうか。マイノリティの保護や、刑事法における手続保障（本書第8章、第9章参照）は、「ひょっとしたら自分もいつかそのような目にあうかもしれない」という立場の交換可能性を前提にした状況がなければ理解されにくい。しかし、本書の見立てに従うのであれば、共感の基礎となる他者との

交流可能性すら寸断、分断されつつある。そうすると、憲法論が持つ規範的なロジックは、今後ますます（この研究会の議論でもしばしばみられたように）「理想を語るものであって現実からは乖離した議論」と見られかねない。

しかし評者は、（あえて砕けた表現をお許しいただければ）身も蓋もない、極めて現実主義的な思考過程と研究アプローチを取る本研究会メンバーとの議論を経て、憲法学が堅持してきた価値論に重きを置く議論スタイルは、今後ますます重要性を増すようにも感じられた。その理由は、研究会の議論において、「結局は人間社会がどのように受け取るか、用いるか」が議論の焦点となる場面が非常に多かったことである。例えば、上述の因果関係が不明な相関関係を示す結果について議論をした際に、「そもそも女性のボーナスを上げるべきか否かは思想と価値の判断の議論であり、今後このような思想と価値判断に関する議論はますます豊潤になるのではないか」という指摘があった。既にAIに関する倫理指針等の議論の盛り上がりからも分かるとおり、人工知能を用いる場面が増大していく過程において、人文・社会科学が果たすべき役割は大きい<sup>(7)</sup>。価値論の体系としても、統治のあり方を示すという観点からも、憲法原理を巡る議論が果たす役割は大きいと考えられる。

また、研究会後改めて本稿を執筆するにあたり再読したところ、本書においてもこの問題意識は明確に示されていることに改めて気がつかされた。それは、第7章の論旨そのものにある。

第7章において、著者（工藤）は近代法・政治システムが基礎に据えてきた個人の自己決定と集団的決定の自律性がフィクションであることを踏まえつつも、「（自然的事実に反していれば制度の正当性が失われるのではなく、逆に）自然的事実とは異なるからこそ、規範的に制度を構築する必要があると言うことも出来よう。そして、現代の法体系では、建前は建前として維持しながら、社会的に適切なバランスを取ろうと様々な形で努力と工夫が重ねられてきた」（338頁）と述べている（傍点は引用者による）。これは、この建前が崩壊してしまえば、制度全体が作り上げてきたバランスをも崩壊させてしまうことへの危機感の表れであろう。そして、第7章自体は、次のように締めくくられる。自律的判断ができる要因がどんどん掘り崩されているか

---

(7) 本研究会後に出版された本として、江間有紗『AI社会の歩き方 人工知能とどう付き合うか』（化学同人、2019年）がある。科学技術社会論（STS）研究者であり世界的な人工知能コミュニティにおける倫理指針策定にも関わる著者が、人工知能と社会の関係について、現時点でもっとも明快かつ横断的な「地図」を示したものとして紹介する。

らといって、それをすべて放棄して、「自己決定は必要無い」という世界（伊藤計劃『ハーモニー』（早川書房、2008年））に移行することは、「非常に魅力的であるが、冒頭のシナリオ以上のディストピアにつながるかもしれず（346頁）、フィクションを維持するための努力の必要性がかつてよりも増している、と（347頁）。

## おわりに

本書の指摘は、効率性や経済合理性を「ほどほどに」統御し、「有用性」との調和をはかるための一里塚である。地方議員のなり手不足問題や社会全体の人手不足問題を踏まえると、AIを使って政策課題を見つけたり、業務の効率化をはかったりする議論そのものは、社会変革を生む好例として取り扱われてきたところである。「地獄への道が善意によって舗装されている」状態をいかに回避するのか、そのバランスの取り方というものが改めて問われることになる。

（よこた あけみ 千葉大学准教授）

キーワード：AI／人工知能／政治過程／政策提言／憲法