

あなたを支配し、社会を破壊する、AI・ビッグデータの罠

キャシー・オニール 久保岡子訳

【目次】

はじめに

Aー・ビッグデータは破壊兵器になる

6

第1章 「モデル」

良いモデル、悪いモデル

26

野球の予測モデル／献立モデル／モデルに埋め込まれた「見解」／再犯モデル——テクノロジーは偏見を排除できない／不透明・規模拡大・有害

第2章 「内幕」

データビジネスの恐るべき真実

53

どのように利益を稼ぐのか／過去の似たパターンを前提に／リスク評価モデルの誤り／現実とは切り離された世界で

第3章 「教育」

大学ランキング評価が多様性を奪う

79

強大な影響力／出口のないラットレース／操作できるランキング／犠牲になる人々

第4章 「宣伝」

弱みにつけこむオンライン広告

107

略奪型のターゲティング広告／相手の「痛点」を探り出す／潜在需要を捉える実験場／リード・ジェネレーション／貧富の格差を増幅する／規制を巧みにかいくぐる

第5章 「正義」

「公平」が「効率」の犠牲になる

130

犯罪を予測する／数理モデルが法の執行を支配する／可能性があるだけでも／公平性か効率性か／刑務所というブラックボックス／地域住民との信頼を築くことから

第6章 「就職」

ふさわしい求職者でも落とされる 160

「適性検査」は適正か？／「個人」ではなく「群れ」を管理する／電子審査と勝ち組／デジタル版のブラインド・オーディションを／才能を発掘するために

第7章 「仕事」

職場を支配する最悪のプログラム 186

勤務時間を効率化するソフト／最低最悪の数学破壊兵器／アイデアが重視される仕事にも／推測のうえに推測を重ねて

第8章 「信用」

どこまでもついて回る格付け評価 212

信用度を格付けする「eスコア」／巻き添え被害／危ない「消費者プロファイル」／システムは「公平性」を判断できない／あらゆるデータが信用データに

第9章 「身体」

行動や健康のデータも利用される 241

群衆から個人へ／カスタマイズサービスの問題点／ドライバーの行動を掌握する／「種族」の生成／健康スコアの波及

第10章 「政治」

民主主義の土台を壊す 268

フェイスブックやグーグルの政治的影響力／有権者個人に合わせてメッセージを送る／政治市場と資金提供者／すべての階層にとって有害に

おわりに
人間だけが未来を創造できる 298

はじめに

AI・ビッグデータは破壊兵器になる

データ経済の弱点

子供のころ、車に乗せられているとき、私はいつも、窓の外を走るほかの車のナンバープレートを見て、数字で遊んでいた。プレートに書かれている数字を、それ以上に小さくは割り切れないところまで分解して遊ぶのだ。たとえば、 $45 \parallel 3 \times 3 \times 5$ というように、すべての自然数は素数をかけ合わせた形で表すことができる。これを「素因数分解」という。私にとつては、探求心を満たしてくれる一番の気晴らしだった。数学マニアの道を歩み始めたばかりだったあの頃から、私は素数に強く興味を引かれていた。

数学に対する私の愛は、やがて情熱へと変わっていく。14歳のときには数学キャンプに参加し、ルービックキューブを胸に抱きしめて帰宅した。数学のおかげで、現実世界の煩雑さから逃れ、整然とした隠れ家に身を隠すことができた。証明に証明を重ねていけば、いくらでも前

に進むことができたし、知識の海原はどこまでも果てしなく広がっていた。しかも、その海原に、私自身の力で知識を追加することもできた。私は大学で数学を専攻し、博士号を取得した。研究テーマは、代数的整数論——私が子供のころに親しんだ素因数分解も、突き詰めればこの分野に入る。そして、コロンビア大学バーナードカレッジの終身在職権付き教授にまでなった。

しかし、私はそこで、人生の舵を大きく切った。教授職を辞し、米国大手ヘッジファンドのD・E・ショーで「クオンツ」〔高度な数学的手法で分析・予測を行う金融工学の専門家〕として働くことにしたのだ。学問の世界から金融の世界に身を移し、それまで抽象的な理論として扱ってきた数学を実践に生かすことにした。私たちが数字に演算を施すと、数兆ドルが口座から口座へと動く。最初は私も、このグローバル経済という新たな「実験室」で、心躍らせたり驚いたりしながら働いていた。ところが、転職から1年あまりが過ぎた2008年の秋、経済は大きな音を立てて崩壊した。

かつて私を保護してくれた数学は、現実世界の問題と深く絡んでいるだけではなかった。数学が問題を大きくしてしまうことも多いという事実を、この経済崩壊はまざまざと見せつけてくれた。住宅危機、大手金融機関の倒産、失業率の上昇——いずれも、魔法の公式を巧みに操る数学者によって助長された。それだけではない。私が心から愛した数学は、壮大な力をもつがゆえに、テクノロジーと結びついてカオスや災難を何倍にも増幅させた。いまや誰もが欠陥が

あつたと認めるようなシステムに、高い効率性と規模拡大性を与えたのも数学だった。

なぜあの時、冷静な頭で考えられなかったのか。経済が崩壊した時点で一步引き返し、なぜ数学が誤った使われ方をしたのか、将来起こりうる同様の惨事を防ぐために何かやれることはないかと考えることもできただろう。しかし、経済危機の後も、私たちは立ち止まらなかった。これまで以上に人々を熱狂させる新たな数学的手法が次々に生み出され、その応用領域は今も拡大し続けている。24時間365日、ペタ（1000兆）バイトもの情報が絶えず粗製乱造（ぞうぞう）されている。その大半は、ソーシャルメディアやeコマースのウェブサイトからかき集められた情報だ。やがて、そのような手法の適用対象は、世界金融市場の動向から私たち人間の動向へと徐々に移っていった。数学者と統計学者が、私たち消費者の欲求、動向、購買力を調査するようになった。消費者の信用度が予測され、学生、就業者としての潜在能力、恋人としての適性、犯罪者になる可能性が計算された。

そう、ビックデータ経済の到来である。これにより、目覚ましい経済発展が見込まれた。コンピュータプログラムを使えば、ほんの1〜2秒のあいだに数千件もの履歴書やローン申込書を処理し、分類し、有望な順に並べた候補者リストを作成することができる。時間の節約になるだけでなく、公正で客観的な資料としてリストを売ることもできる。偏見をもつ人間が大量の書類に目を通すのではなく、ただの機械が血の通わない数字を淡々と処理するのだから。2010年ごろには、人事部門での数学の存在感はこれまでになく高まり、大きな期待をもつ

て迎え入れられた。

でも、私には弱点が見えていた。数学の力で動くアプリケーションがデータ経済を動かすといつても、そのアプリケーションは、人間の選択のうえに築き上げられている。そして、人間は過ちを犯す生き物だ。モデルを作成する際、作り手は、最善の意図を込め、良かれと思って選択を重ねたかもしれない。それでもやはり、作り手の先入観、誤解、バイアス（偏見）はソフトウェアのコードに入り込むものだ。そうやって作られたソフトウェアシステムで、私たちの生活は管理されつつある。神々と同じで、こうした数理モデルは実体が見えにくい。どのような仕組みで動いているのかは、この分野の最高指導者に相当する人々——数学者やコンピュータサイエンティスト——にしかわからない。モデルによって審判が下されれば、たとえそれが誤りであろうと有害であろうと、私たちは抵抗することも抗議することもできない。しかも、そのような審判には、貧しい者や社会で虐げられている者を罰し、豊かな者をより豊かにするような傾向がある。

アルゴリズムが人間を選別する

私は、そのような有害なモデルを「数学破壊兵器（Weapons of Math Destruction : WMD）」と呼ぶことにした〔大量破壊兵器のMassと数学のMathをかけている〕。数学の何がどのように「破壊

兵器」なのか、ある事例を紹介しながら説明しよう。

よくあることだが、これから取り上げる事例も、最初に掲げられた目標自体は称賛に値するものだった。2007年、新たにワシントンDC市長になったエイドリアン・フェンティは、学力水準の低い市内の学校の状況改善を決意し、身を粉にして働いた。当時、第9学年を修了して卒業まで漕ぎつける高校生は2人に1人、第8学年で数学の成績が学年水準に達している学生はわずか8%というあり様だった。フェンティは、新たにワシントンDC学区の教育総長という役職を設置して強い権限をもたせ、教育改革論者のミシェル・リーを起用した。学生の成績が十分でないのは教師の教え方が悪いからだ、という理論に立ち、2009年、リーは業績の低い教師を一掃する計画を実行に移した。この施策は、問題を抱える全米各地の学区で導入が検討されており、システムエンジニアリングの観点で言えば、完全に筋の通った考え方である。教師を評価し、評価の低い者を除外し、評価の高い者を、その力が最大限に発揮される場所に配置する。データサイエンティストの言葉でいうなら、これは、学校システムの「最適化」であり、子供たちにとって、より良い結果を生むものと思われた。「無能」な教師を別にすれば、誰からも文句は出ないはずだ。リーは、IMPACTと呼ばれる教師評価ツールを考案し、2009〜10年の学年度末に、評価スコアが学区内の下位2%に入る教師を全員解雇した。翌年にはさらに、下位5%に相当する206名が解雇された。

第5学年の教師だったサラ・ウィソッキーは、自分は何の心配もないと思っていた。マク

ファアランド中学での教師歴はまだ2年だったが、すでに生徒からも保護者からも「優秀」な教師として評価されていた。なかでも、子供たちに対するきめ細やかな配慮は高く評価されており、「これまでにお世話になった先生のなかで最も素晴らしい」という声も聞かれるほどだった。

ところが、2010〜11年の学年度末、ウィソッキーのIMPACTによる評価スコアは散々な結果だった。教師の数学と国語（英語）を教える能力を測定する目的で新たに導入された「付加価値モデル」での評価スコアが著しく低かったのだ。このスコアは、アルゴリズムによつて算出されるもので、評価全体の半分を占めており、学校経営陣や地域コミュニティから受けた高い評価も、その穴埋めにはならなかった。学区としてはほかに選択の余地はなく、ウィソッキーは、IMPACTの評価スコアが最低水準に達しなかったほかの205名の教師とともに解雇された。

彼女は、いわゆる魔女狩りに遭ったわけでも、不正にスコアを低くされたわけでもない。学区の側には、それ相応の論理があった。学校経営陣も人間なので、どうしても、無能な教師と馴れ合いの関係になることがある。その教師の流儀や見かけ上の熱心さを高く評価することもある。無能な教師が良い先生のように見えることもある。だからこそ、ワシントンDC学区は、ほかの多くの地区と同様、人間のバイアスを最小限に抑え、確固たる結果に基づくスコアをより重視することにしたのだ。

もちろん、ウィソッキーは自分のスコアを見て、ひどく不当だと感じ、なぜこのような結果になるのかを知りたがった。しかし、「なぜそうなるのかを理解できている人なんて、1人もいないように思います」と、後日、彼女は私に語った。優秀な教師のスコアが、どうしてこんなにも低い点数になるのか？ 付加価値モデルはいったい何を測定しているのか？

彼女が調べてわかったことは、複雑な評価が行われていて、とても理解できそうにないということだった。ワシントンDCは、プリンストンに拠点のある政策数理研究所 (Mathematica Policy Research : MPR) とコンサルタント契約を結び、評価システムを考案した。MPR が取り組んだ課題は、学区内の学生の学習進捗度を測定したうえで、学力の向上・低下に対する担当教師の貢献度を算出することだった。もちろん、簡単にできることではない。学生の成績には、その学生の社会的・経済的な背景から学習障害の影響に至るまで、数多くの変数が関係するということは、MPR の研究員も知っていた。アルゴリズムを作成する際には、そのような状況の違いを割り引く仕組みが必要であり、それも、複雑化の一因になった。

実際、人間の行動や能力、潜在能力をアルゴリズムに落とし込むというのは、至難の業だ。MPR が直面した課題について理解を深めるために、ワシントンDC南東部の貧困地域に住む10歳の女の子について考えてみよう。ある年の学年度末に、その女の子は第5学年の標準学力テストを受ける。そして、その後も彼女の生活は続く。家族の問題や、お金の問題を抱えるかもしれない。近いうちに引越すことになるかもしれない。あるいは、法的なトラブルに

巻き込まれた兄のことが心配でたまらない状況になるかもしれない。自分の体重のことで悩んだり、学校でいじめに遭って苦しんだりする可能性もある。いずれにせよ、翌年、彼女は第6学年の標準学力テストを受ける。

今回と前回の学力テストの結果を比較した場合に、スコアは安定しているか、できれば急上昇してほしいところだ。彼女の結果が下がっていた場合、彼女の成績と他の成績優秀者の差を計算するのは簡単である。

しかし、その差のうち、どの程度までが教師の責任になるのか？ これは難しい問題だ。MPR のモデルでは、ほんのわずかな数字の比較しか行われていない。グーグルのようなビッグデータ企業であれば、研究員が絶えず試験を繰り返し、何千個もの変数をモニタリングするところだ。グーグルでは、広告を1つ表示するにも、テーマ色を赤と青のどちらにするかを決める前に、各色のバージョンの広告を数千人に提示し、どちらがより多くクリックされるかを追跡することができる。そうやって得られたフィードバックを自社のアルゴリズムに反映し、広告表示の運用を細かく調整して精度を高めている。グーグルのやり方にも問題は山ほどあるし、追々、取り上げるつもりだが、いずれにしても、グーグルの試験方法では、統計学が有効に活用されている。

学年度を通じて1人の人間が他者に与える影響の大きさを計算しようと試みるのは、これより遥かに複雑になる。「学ぶ側にも教える側にも多くの要因がかかわってくるため、そのすべ

てを測定するのはきわめて困難です」とウィソツキーは言う。さらに言えば、わずか25〜30名の学生のテスト結果だけを分析して教師の技量をスコア化しようとすること自体、馬鹿げている。少なすぎる数のサンプルで全体を測ろうとすれば、おかしい結果が出る可能性も高くなる。現に、検索エンジン並みの統計学的な厳密さで教師を分析しようと思ったら、無作為に選出した学生を数千〜数百万回は試験しなければならぬ。まともな統計学者は、外れ値や異常値を相殺させるために、膨大な数のサンプルを集める（一方、数学破壊兵器となるモデルでは、後述のとおり、たまたま外れ値となった個人が不当に扱われがちである）。

もう1つ、サンプル数と同じくらい重要なポイントがある。統計システムには、フィードバックがなければならぬ。何かが標準から逸脱していたなら、そのことを明示的に返す必要がある。統計学者はエラーを利用して自分たちのモデルを育て、磨いていく。たとえばAmazon.comの「おすすめ」は、最初のうちは相関が不完全で、10代の少女に芝生の手入れに関する本をすすめるようなことをしていたが、クリック数が極端に少なければ、クリック率が伸びるまでアルゴリズムの微調整が繰り返される。しかし、フィードバックがなければ、統計エンジンはいつまでも見当違いな商品をおすすめし続け、誤りから学習することなく、有害な分析を続けることになりかねない。

本書で論じる数学破壊兵器の多くは、ワシントンDC学区の付加価値モデルも含めて、フィードバックがないまま有害な分析を続けている。自分勝手に「事実」を規定し、その事実

を利用して、自分の出した結果を正当化する。このようなたぐいのモデルは自己永続的であり、きわめて破壊的だ。そして、そのようなモデルが世間にはあふれている。

政策数理研究所(MPR)のスコアリングシステムがサラ・ウィソツキーと他の205名の教師をはじめ出したとき、学区側は、直ちに彼らを解雇した。では、システムの評価が正しかったかどうかを、システムはいつ学習するのか？ 学習はしない。システムが「失格」と判定すれば、それがその教師の評価になる。そして、206名の「無能」な教師がいなくなつた——その事実だけで、付加価値モデルの有効性が実証されたかのようにみなされている。システムによる評価の低い教師は学区から一掃される。真実が調査されることのないまま、スコアによって現実が作られていくのだ。

人生をひっくり返す

これは、数学破壊兵器のフィードバックループの一例にすぎない。本書では、ほかにも多くの事例を紹介していく。たとえば、最近では、従業員を採用する際に、クレジット(信用度)スコアを利用して候補者を評価する雇用主が徐々に増えている。請求書の支払いを滞納したことのない人物なら、遅刻せずに出社し、社内ルールを順守する可能性が高いと考えてのことだ。現実には、責任感の強い優良な働き手であっても、ちよつとした不運で苦境に陥り、クレ

ジットスコアを落とした人が大勢いる。しかし、クレジットスコアの低さと職能レベルの低さには相関があるという考え方のせいで、スコアの低い人は職に就きにくくなっている。無職の状態が続けば貧しくなり、貧しい状況が続けばクレジットスコアはさらに悪くなり、ますます就職は困難になる。まったくの悪循環だ。雇用主の側にしても、クレジットスコアを重視したせいで、いったいどれほどの数の優良な従業員を逃してきたのか、知る方法がない。数学破壊兵器には、たぐさんの有毒な「思い込み」が、数学によって偽装された状態で搭載されており、検証されることも疑念を抱かれることもないまま、広く世に出回る。

こうした事例は、数学破壊兵器のもう一つの大きな特徴を浮き彫りにしている。このたぐいのモデルには、貧しい者が不当に扱われる傾向があるのだ。その一因は、大勢の人を評価する目的で設計されている点にある。大量処理に特化していて、価格も安い。それが売りでもある。一方で、裕福な人は、個人情報インプットによって利益を得ることが多い。アイビリーグ出身者が多く働く大手の法律事務所や富裕層の多い排他的な名門私立校は、ファーストフードのチェーン店や財政難にあえぐ都市部の学区に比べると、紹介者からの推薦や対面での面接を重視することが圧倒的に多い。この後、本書で何度も確認することになるが、一般に、特権階級の人ほど対面で評価され、庶民は機械的に評価される。

ウィソッキーは、スコアがあそこまで低かった理由を誰にも説明してもらえなかった。この事実からもわかるように、数学破壊兵器が出した評決は、アルゴリズムの神々からの預言のよ

うに天から降って来る。モデル自体はブラックボックスで、中身は企業機密として厳重に保護される。おかげで、政策数理研究所のようなコンサルタント業者は、高い代金を請求できるし、同じモデルを別の目的に使うこともできる。評価される人々には何も知らされないのです、評価される人々がシステムをうまく利用しようと試みる可能性も低い。となると、人々はただ懸命に働き、ルールを守り、彼らの努力をモデルが正しく記録して適切に評価してくれるように祈ることしかできない。詳細が隠されたままでは、スコアに疑問を呈することも抗議することも一層難しくなる。

不当なスコアの理由を知ろうとしたのはウィソッキーだけではなかった。何年ものあいだ、ワシントンの教師たちは独裁的なスコア制度に不満を抱き、モデル内部で行われた計算の詳細を求めた。しかし、アルゴリズムによって算出される、としか説明されなかった。中身はきわめて複雑なのだ、と。この説明に多くの人は氣力を削がれ、それ以上の説明を要求しなくなる。残念なことに、多くの人が数学に怯え、言いなりになっているのだ。だが、引き下がらなかった教師もいる。数学教師のサラ・バックスは、学区の役員をしているジェイソン・カムラスと元同僚でもあったため、彼に詳細を説明するよう強く要求し続けた。数カ月間にわたる押し問答の末、カムラスは彼女に、もうすぐテクニカルレポートが届くから待つように、と告げた。これに対してバックスは、「自分で説明もできないような方法で人々を評価することを、どう正当化するつもりですか？」と返した。しかし、数学破壊兵器とはそういうものののだ。

分析は外部のプログラマーと統計学者に委託される。また、原則として、説明はマシンに任せられる。

一方のサラ・ウィソッキーは、説明は受けられなかったものの、生徒の標準学力テストの点数がアルゴリズムの公式のなかで重みづけされていることには気づいていた。だとしても、まだ疑念は残った。マクファーランド中学で過ごした最後の年が始まる前、彼女は、自分がこれから受け持つことになる生徒について、前年度末の小学校の学力テストの点数が驚くほど良いことに気づき、喜んでいたので。サラが受け持つ生徒の大半は、バーナード小学校の出身だった。バーナード小学校では、学生の29%が「上級レベルの読む力」と評価されていた。これは、この学区の平均の5倍に相当する。

ところが、実際に授業を始めてみると、生徒の多くは簡単な文を読むのにも苦勞していた。すいぶん後になって、ワシントン・ポスト紙やUSAトゥデイ紙の記者が明らかにしたところによると、この学区内の41校で、標準学力テストの回答用紙に答えを消した跡が大量に見つかり、その41校のなかにバーナード小学校も含まれていた。高い率で答えが修正されていたということは、不正行為が行われた可能性が高い。学校によっては、70%の教室で不正が疑われた。

この不正行為に、数学破壊兵器の何が関係しているのか？ いくつか考えられるが、第一に、教師評価アルゴリズムは教師の行動に大きな影響を与える強力なツールである。もともとそのために開発されたツールであり、実際にワシントンの学校では、教師にとつて^{アメ}餌にも^{ムチ}鞭に

もなった。教師は、生徒が学力テストで失敗すれば自分の職が危うくなると知っていた。それが、生徒にテストで確実に良い点数を取らせたいという教師の動機付けになった。世界的な金融不況で労働者市場が大打撃を受けていた時期なら、なおさらだ。しかも、自分の生徒の成績がほかの生徒よりも優秀であれば、教師と学校経営陣は最高8000ドルの特別手当をもらえる。今回のように、大量の消し跡と異様に高い点数という証拠に、このような強力な動機が加われば、小学校の教師が不安や欲に屈して生徒の回答を修正したのではないかと疑うのに十分な根拠となる。

となると、サラ・ウィソッキーが受け持った生徒の学年開始時のスコアは人為的に水増しされていたと考えられる。だとすれば、彼らの学年度末の学力テスト結果を評価する土台は失われており、担当した教師の評価は実際よりも低くなるだろう。ウィソッキーは、自分の身に何が起こったのかに思い至り、確信を深めた。この説明なら、保護者や同僚、そして生徒たちから受けた「優秀な先生」という見解ともつじつまが合う。彼女を困惑させた疑念も解決される。サラ・ウィソッキーは明確な反論の証拠を得た。

ところが、数学破壊兵器に対しては、訴えかけることはできない。そこが、この破壊兵器の恐ろしいところでもある。聞く耳をもたないのだ。一度出した評決を曲げることはない。甘い言葉になびかず、脅してもおだても聞く耳をもたないのはいいが、論理的に説明しても受け入れられないのは困ったものだ。結論として出力されたデータに対して正当な疑問がある場合

でさえ、受け入れられることはない。もちろん、自動化されたシステムが厄介な問題を含んだ基準のせいで失態を演じた事実が明らかになれば、プログラマーが過去に遡ってアルゴリズムを微調整してくれるだろう。しかし、大方の部分については、プログラムはひるむことなく評決を下し、プログラムを採用している人間は肩をすくめることくらいしかできない。「きみに何ができるんだい？」と言わんばかりだ。

サラ・ウィソツキーが学区側から最終的に受けた対応は、まさにそうだった。後に、ジェイソン・カムラスはワシントン・ポスト紙の取材に対し、こう語っている——「回答用紙の消し跡は『示唆的』であり、彼女が受け持った第5学年のクラスの数値には誤りがあった可能性はある。しかし、決定的な証拠ではない。彼女に対する処遇は公正だった」

これはもう、完全なバラドックスだ。アルゴリズムによって大量の統計処理が施され、ある人物が、採用しないほうがいい候補者、リスクの高い借り手、テロリスト、無能な教師である可能性が算出される。この「可能性」はスコアに変換される。そして、時に、そのスコアが1人の人間の人生をひっくり返す。しかも、人生をひっくり返された人物が反撃に出ても、「示唆的」な証拠だけではどうにもならない。この事例の場合、事態を覆すのは厳しに違いない。数学破壊兵器の犠牲者をこれから何人も紹介していくことになるが、犠牲者はみな、アルゴリズムそのものに要求されるよりも遥かに高い水準で証拠を求められる。

サラ・ウィソツキーは、解雇されてショックを受けたものの、失業期間はほんの数日だっ

た。彼女には味方が大勢いたからだ。彼女の生徒たちも、彼女が優秀な教師であることを保証してくれた。そしてすぐに、バージニア州北部の裕福な地域にある学校に就職することができた。大いに疑わしいモデルに頼ったせいで、貧しい学校は優秀な教師を失い、生徒の学力テストの点数に頼って人を解雇したりしない裕福な学校は、優秀な教師を迎え入れた。

有害で悪質なフィードバックループ

住宅危機のあと、私は、数学破壊兵器が銀行業界に蔓延していることや、銀行が私たちの経済を危険に晒しているという事実が目覚めた。2011年前半には、私はヘッジファンドでの仕事を辞めた。その後、データサイエンティストとしての自分のブランドを見つめ直し、eコマースのスタートアップ事業に参加した。そのポジションから多くのことを見聞きしたおかげで、私は、ほかにも数学破壊兵器が大量に存在し、ありとあらゆる業界で人々を攪乱していること、その多くが不平等を増幅させ、貧しい者を罰していることに気づくことができた。数学破壊兵器は、勢いを増すデータ経済の中心部に潜んでいる。

数学破壊兵器について広く知ってもらうために、私は「マス・ベイブ (mathabe)」というブログを開設した。私の狙いは、数学者仲間を集め、自己永続的で有害なフィードバックループを生み出す中途半端な統計学や偏見に満ちたモデルの使用に反対することだった。とくに

データサイエンティストからは高い関心が寄せられた。ブログに集まったメンバーから、数学破壊兵器が新たな分野に次々に広がっているという警報も受け取るようになった。しかし、2011年の半ば、「ウォール街を占拠せよ」を合言葉にマンハッタン南端部で抗議運動が発生して勢いづくころ、私は、自分にはより広い公共の場でやるべきことがあるのだと気づかされた。経済的公正さと説明責任を求めて、何千人もの人々が集まっていたのに、ウォール街を占拠した人々がインタビュアーに答える様子を見てみると、多くの人は、金融に関する基本的な問題について知識を持ち合わせていないようだった。彼らが私のブログを読んでいないのは明らかだ（とはいえ、もちろん、システムがうまく機能していないことを知るために、システムの詳細をすべて理解している必要のないことは、追記しておかなければならない）。

私は、ウォール街の抗議運動に対して、批判することも参加することもできるのだと気づき、参加した。間もなく、コロンビア大学で週1回開催されるオルタナティブ銀行グループのミーティングで進行役も務めるようになった。ミーティングでは、金融改革について議論が交わされた。このプロセスを通じて、私はあることに気づいた。大学を飛び出し、金融とデータサイエンスという2つの分野のベンチャーを経験したおかげで、私はいつの間にか、数学破壊兵器に力を与えているテクノロジと文化の両方について驚くほど詳しくなっていた。

現在、ビジネスから刑務所まで、社会経済のあらゆる局面の細かな管理が、設計に不備のある数理モデルによって遂行されている。それらの数学破壊兵器には、ワシントンの公立学校で

サラ・ウィソツキーのキャリアを狂わせた付加価値モデルと同じ特徴が多く備わっている。中身が不透明で、一切の疑念を許さず、説明責任を負わない。規模を拡大して運用されており、何百万人もの人々を対象に、選別し、標的を絞り、「最適化」するために使用されている。算出された結果と地に足の着いた現実とを混同することによって、有害で悪質な数学破壊兵器のフィードバックループを生み出しているものも多い。

ただし、学区で採用された付加価値モデルと、法外な高金利で金を貸すペイデイローンの見込み客を探し出す数学破壊兵器とのあいだには、1つ重要な違いがある。この2つは、「見返り」が異なるのだ。学区の付加価値モデルの「見返り」は、ある種の政治の世界の通貨、つまり問題が改善されつつあるという感覚である。一方、ビジネスでモデルを使用する場合の見返りは、通常の通貨、つまりお金である。人をだますためにアルゴリズムを活用するようなビジネスの場合、金が流れ込んで入れば、モデルが機能していることの証明になる。モデル制作者の目を通して見れば、筋が通っている。顧客を見つけ出す目的や絶望的な借り手を巧みに操る目的で統計システムを構築する場合は、利益が膨らんでいけば、アルゴリズムは正しい手順を踏んでいるとみなされる。ソフトウェアは自分の任務を遂行している。問題は、やがて利益率が真実に取って代わるようになる点である。この危険な「事実の混同」は、本書でも繰り返し登場する。

そのような「混同」が起きるのは、データサイエンティストが、一連のデータ処理の影響を

受ける側にいる人々の存在を簡単に見失うからである。データを高速処理するプログラムは、人々について、ある一定の割合で誤解することを避けられない。人々を誤ったグループに分類し、就職の機会や夢のマイホームを建てるチャンスを奪うこともあるのは、もちろん、データサイエンティストも理解している。しかし、原則として、数学破壊兵器を運用する人々は、そのようなエラーについてあまり深く考えない。彼らにとつてのフィードバックはお金であり、それがインセンティブにもなっている。彼らが作るシステムは、より大量のデータを処理し、より多くのお金流れ込んでくるように分析を微調整するように設計されている。もちろん投資家も、その見返りを享受し、数学破壊兵器を売る会社に対し、より多くのお金を投じるようになる。

では、数学破壊兵器の犠牲者はどうなるのか？ 社内のデータサイエンティストは、完璧な統計システムなど存在しない、と言うだろう。巻き添えになつて損害を被る人がいるのは仕方がない、と。そして犠牲者は、サラ・ウィソッキーのように、取るに足らない人間、失つてもかまわない人間のように扱われる。犠牲者のことは忘れて、おすすめ商品を推薦するエンジンのおかげで役立つ情報を得た人や、無料動画サイト「パンドラTV」で大好きな音楽を見つめることのできた人に目を向けよう、と彼らは言うかもしれない。圧倒的大多数の話に目を向け、小さな傷は無視しよう。

AI・ビッグデータには、大勢の熱烈な支持者がいる。だが、私は違う。本書では、世間の流れに逆らい、数学破壊兵器によつて生じた損害や数学破壊兵器によつて延々と生み出される不正行為に注目し、鋭く切り込んでいく。大学への進学、お金の借入れ、刑務所行きの判決、職探しや昇進など、人生の重要な瞬間に有害な影響を受けた人々の事例を数多く見ていくことになる。あらゆる生活領域で、独裁的に罰を振りかざす「秘密のモデル」による支配が進みつつあることに、あなたも気づくだろう。

それでは、AI・ビッグデータの暗黒面を覗いてみよう。

データサイエンス

AI・ビッグデータの活用は、いまや誰もが被害に遭うかもしれない危険水域に入っている。にもかかわらず、その実態は余りにも知られていない。こうしたダークサイドは、よく耳にするAIが人間の仕事を奪うといった危惧などとはまったく異なる。「人間の仕事を奪う」という予測は、むしろAIの素晴らしさを認めている裏返しなのだから。本書が明らかにするのは、AI・ビッグデータの仕組みや活用法そのものの中に、私たちの人生や社会を狂わせ、壊すようなリスク（罠）が潜んでいるということだ。データビジネスの現場を熟知するデータサイエンティストである著者は、具体例をあげながら、こうした実情をわかりやすく伝えてくれる。

読者は、「教育・宣伝・仕事・信用・健康・政治・犯罪」……といった日常の多くの場面で、データ活用が進んでいることを知るだろう。その領域は、従来のAI・ビッグデータ活用の範囲を踏み越えて広がっている。私たちの能力・適性・信用、さらには善悪や身体といった、そもそも定量化し数値に置き換えるのが難しいものごとまで、データとして計算され、評価されているのだ。

本書で強調されるように、AI・ビッグデータを動かすアルゴリズムはけっして万能ではなく、むしろ欠陥だらけのまま使われている。たとえば、アルゴリズムの作り手の先入観や誤解、偏見などが無意識のうちに紛れ込む。過去のパターンを参照するため、従来の価値観が踏襲される。

また、元になるモデルは、扱う事柄を単純化したもので、リアルな世界の複雑さや機微などは、こぼれ落ちてしまう。重要な情報も抜け落ちる。さらに、どのようなデータを集め、どんな目的に使うのかという選択にも、作り手や企業の価値観やニーズが反映される。つまり、アルゴリズムとは、数学のなかに「見解」や「成功の定義」を埋め込んだもののなのだ。

本書は、とくに悪質なアルゴリズムを、「数学破壊兵器」と呼ぶ。その要素は、仕組みが「不透明」で修正されにくく、急速に成長する「規模拡大」性をそなえ、不公平な内容などによって「有害」な影響を及ぼす——という3点である。その影響力が強大になると、評価される側は、アルゴリズムに気に入られるように、行動・態度を変えらるることさえある。単純化された同じ指標を満たさざるをえなくなり、多様性が失われていくのだ。また、職場での能力も、数値化されない要素（場を和ませるコミュニケーションなど）も大きいのに、こうした人間的スキルは評価されようがない。

こうしてアルゴリズムに評価されなくなると、そこに有害なフィードバックループが生まれる。たとえば、信用格付けが落ちれば、就職やローンなどで不利になり、その影響がさらに他分野に及び……と悪循環にはまり、格差が広がっていく。データの誤情報などで、被害を受けることもある。そんな場合でも、いったいなぜ、自分の評価が落とされたり、間違つて類型化されたかなどは、通常わからない。アルゴリズムの中身は、企業機密として厳重に保護されるなどしているからだ。近年では、さらに進んで、個人の行動や身体までもが、データ化され、評価・選別されるようになってきている。たとえば、企業が従業員の健康データを収集し、報酬に結びつけはじめている。

選挙では、有権者個人にターゲットを絞った「マイクロターゲティング」が盛んだ。このような有権者個人を選別するシステムは、一部の有権者の重要度を高めることで、民主主義の土台を壊すことにつながる。また、消費者のプロファイリングは、分類された「種族^{トライブ}」として階級化を促すとともに、私たちの欲求を枠付け、コントロールしていくようになる。

本書は主に米国のデータ活用の事例を扱っており、その社会事情（個人情報の扱いや人種差別など）とも関係している。だが、本書が注目するのは、たんなる事例を超えた、データ活用の中核となる仕組みや目的なのだ。もともと、データ活用は、人間が陥りやすい偏見や誤謬を避けて、公平・中立な判断ができるという善意に基づくことも多い。しかし、すでに明らかにになったように、データにも偏見・誤謬はつきものだ。それでも、公平・公正よりも、効率・利便性・収益が優先される。

ある分野での規制を厳しくしても、分野を超えて次々と増殖していくのが、アルゴリズムの特性だ。個人情報の取り扱いを制限したところで、規制をかいぐり、代理データなどで目的を果たす。そう、人間を選別し、作業を効率化し、収益を上げるといった「目的」のもとにある限り、いたるところで有害な破壊兵器は増殖していく。本書は、けつしてAI・ビッグデータの活用そのものを否定しているわけではない。有害なアルゴリズムであっても、使い方次第で生まれ変わる。そのために、私たちはアルゴリズムを監査するなどして、主導権を人間に取り戻さなくてはならない。モラルのある想像力によって、未来を創造できるのは、人間だけなのだから。