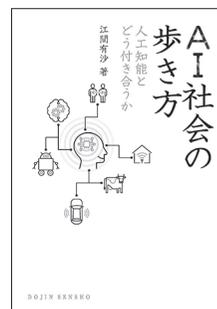


## 書評01

江間 有沙 著

『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2月刊 / 272ページ / 2000円+税  
ISBN 978-4-759-81680-8評者：上野 秀剛  
奈良工業高等専門学校

本書は科学技術社会論がご専門の江間有沙先生による、人工知能（AI）と社会についての優れたまとめです。この分野は科学技術が社会にもたらす影響の分析や課題の指摘、対策を講じる分野です。本書ではAIの技術以外の要素である制度設計や研究者・開発者の倫理、社会のあり方、利用者の関わり方について様々な取り組みと実践者の主張、各種団体による指針などを紹介しています。また、「現時点の技術で何ができるか」という観点よりも、未来の技術を導入した、未来の社会はどうあるべきか、どのようなプロセスで進んでいく（べき）か、という点に焦点が当てられています。

1960年代に始まった第1次AIブームでは、専門家の知識を人間がルールとして記述し、そのルールに従ってデータを処理するプログラムがAIと呼ばれました。現在の第3次AIブームでは多量のデータ（ビッグデータ）から確率・統計に基づいて傾向を自動で学習する「機械学習」が用いられます。例えば動物の写真を大量に集め、その種類が区別できる特徴（例えば顔の形や体の模様）を人間が指定することで、AIは動物ごとの特徴の違いを（統計的に）学習します。この結果を基に、初めて見た画像の特徴と、過去のデータの類似性からどの動物か予測します。深層学習（Deep learning）と呼ばれる手法では特徴が現れる箇所もAIが自己学習するため、人間による指定が不要です。AIは各種装置に内蔵されたセンサーのデータ

を使った故障時期の予測や、生活習慣・活動データ（スマートウォッチで計測する脈拍など）を用いた健康リスク評価など幅広い分野で既に使われています。多くの場合、開発者が用意したデータで学習済みのAIが装置・サービスに組み込まれ、利用されます。しかし、データを統計的に処理する、という特徴のため、例えばある特定のグループに属する人たちのデータが学習に用いられず、信用評価AIに不当に低い評価をされてお金を借りられない、といったトラブルも起こりえます。また、同じデータであってもAIの学習方法（アルゴリズム）が異なれば、判断に用いる特徴やその重要度が異なるため、データとアルゴリズムの双方において不公平・不平等が起きない仕組みが必要です。AIは各現場のデータを追加で学習することもあるため、利用者は学習・運用が正しく行われているか常に注意を払う必要があります。また、開発者や社会がもつ無意識のバイアスによって学習データや使用アルゴリズムの選択がゆがめられ、AIに組み込まれる恐れもあります。意図的にゆがんだデータを学習させることでAIの判断をゆがめる「攻撃」も存在します。このような事象に対する対策は技術者だけでは無く、データが使われる立場の利用者も巻き込んだ社会全体での議論と合意、判断が必要です。特に、（現時点の）AI技術の多くはなぜそんな判断になったのか説明できないため、AIによる判断が本当に妥当・公平なのか、説明可能な人工知

能 (eXplainable AI: XAI) の開発が必要とされています。説明可能だけでなく、人がその説明を受け入れる必要もあります。例えば裁判システムが 0.1 秒で出した判決は、たとえ幅広いデータに基づいた妥当なものであっても多くの人間には受け入れられないと思われます。AI が説明可能な「理由」を示すとともに、その提示方法 (ユーザインタフェース) の開発や人間自体の理解も重要です。人間の心を理解することで (例えばロボットペットに対して愛情を抱くように) 信頼してもらえる AI が作れるかもしれません。AI を用いたシステムの公平性、アカウントビリティ、透明性についての研究 (各単語の頭文字をとって FAT と呼ばれます) が近年注目を集めており、単に高精度なだけでは無い「有益な」人工知能の開発が待たれます。また、倫理的な指針やガイドラインの整備も進んでおり、研究者や開発者が守るべき「研究倫理」、技術の運用に関する「人工知能倫理」に加え、倫理的な価値観を組み込んだ「倫理的な人工知能」についての指針の整理が進みつつあります。倫理的な判断はトロツク問題のように人間でも難しい場合がありますが、自動運転を行う AI には近々の課題でもあります。今後、どんな倫理をどのように AI に組み込むか社会全体で考える必要があります。さらに社会的な対応としては、自分のデータが使われる範囲を自分で決定、制御できる仕組みである情報銀行についても指針が公開されています。

AI が社会に与える影響としてよく話題に上がる「機械に仕事を奪われる」についても本書では触れています。著者は「短期的 (長くても 10 年程度) には人間の仕事は奪われない、ただし仕事のタスクが変わる」という見方を紹介しています。我々は既に仕事の一部 (タスク) を機械に任せています。これと似たようなタスクを AI に奪ってもらうことに抵抗を持つ人は少なく、それ以外のタスクが仕事としては残ります。一方で、人と人が接する仕事における「接

客タスク」は労働者と顧客双方の尊厳や責任などの社会的価値と結びついているため、たとえ AI に可能であっても拒む人も多いでしょう。機械化が受け入れられやすい部分 (かつ、技術的に機械化しやすい部分) と、人間がやることに価値がある部分を区別しながらどんなビジネスモデルを組み立てるのか、経営者は考える必要があります。また、労働者としての人間に求められるのは、ルール化できない (機械に任せられない) 例外的な処理への対応力と、機械に合わせて仕事の仕方を変える適用力であると筆者は指摘しています。

評者は情報系の研究・教育を仕事にしており、研究で AI の作成を日常的にしていますが、非専門家の方に AI ができること・できないことを説明することの難しさを日々感じています。今後、専門家以外の方が AI を安心して使うためには、自動車を取り巻く法制度や、教育体制、技術的な点を信託できる車検制度と同じような、仕組み作りが重要であると再認識させられました。また、新しい技術が流行ると、組織への導入自体が目的化することも珍しくありません。特に効率化が重視されがちな昨今では当たり前のように AI で効率化! という考えになってしまいます。しかし、本書にも書かれているように、技術は利用者を幸福にするためのものです。AI 技術は社会に大きなインパクトを与えることは間違いありません。AI を導入することでより幸福な暮らし (と、その一部としての働き) ができるような社会システムの構築とそのための議論、そしてそのための社会の理解が今後ますます重要になります。人間、文化、制度、そして技術の複合物である社会を理解できたとき、手触りを感じてその存在に接することができる、理想の人工知能の姿と、その付き合い方が見えてくるのではないのでしょうか? とここで、付き合い方といえば、AI に法人格を与える (AI が法的義務や人権を持つ) ことについて、皆さんどう思われますか?