

# A I、データサイエンスで日本をどう変えていくか

東京大学大学院工学系研究科  
人工物工学研究センター／技術経営戦略学専攻 教授  
松 尾 豊

## ≪構 成≫

- 1 人工知能の現状
- 2 我が国の人工知能関係の動き
- 3 松尾研究室を中心とする活動
- 4 今後の処方箋

### 1 人工知能の現状

人工知能 (Artificial Intelligence; A I) の技術が進展している。中でも、深層学習 (ディープラーニング) という技術は、2010 年代に大きく躍進し、第 3 次人工知能ブームの中心となっている。深層学習は、2012 年に画像認識における大きな性能の飛躍で注目された。2016 年に、アルファ碁がイ・セドル 9 段を破ったことは記憶に新しい。アルファ碁は深層強化学習という技術が用いられており、英国ロンドンの A I 企業であり、Google の子会社である DeepMind が開発した。その後も、深層学習は着々と進展を遂げ、最近では、自然言語処理において、「トランスフォーマー」という機構が席卷している。トランスフォーマーは、自己アテンションという仕組みを巧妙に使ったニューラルネットワークの新しいタイプであり [1]、様々な処理を内部に学習することができる。2020 年に公開された G P T - 3 というモデルは、1,700 億のパラメータを持ち、5 兆単語で学習されており、様々な自然言語

処理のタスクに非常に高い性能を発揮する [2]。翻訳や質問応用、要約などはもちろんのこと、表を作成する際に自動的に情報を補足するようなデモ<sup>1</sup>やプログラムを自動で生成するようなデモ<sup>2</sup>まで公開された。シリコンバレーにある OpenAI という非営利の研究組織が作ったものだが、早速マイクロソフトが独占的なライセンスを受けた。いずれ、マイクロソフトの製品に組み込まれていくことになるのだろう<sup>3</sup>。

今年の 7 月には、AlphaFold 2 という技術が公開された [3]。アルファ碁を開発した DeepMind が出したものであり、タンパク質の 3 次元構造を推定するというものである。この問題は古くから難問として知られているが、深層学習を活用することで、従来の精度を大きく超え、非常に正確な推定を可能にした。バイオテクノロジーの研究者から、何年も解けなかった問題が解けたと絶賛されている<sup>4</sup>。深層学習は、今後、材料開発や医薬品の開発など、科学技術の基礎に関わる部分でも大きな影響を与えていく可能性がある。

ほかにも、深層学習の実用化は世界で相次いでいる。例えば、医用画像における画像診断は、米国や中国で次々と実用化した。画像や映像からのセキュリティのシステムにおいても、深層学習は使われている。米国や中国

<sup>1</sup> 例えば、東京 1,351 万、神奈川 912 万と入れただけで、人口についての表を作っていると解釈し、大阪と入れたとたん、883 万と自動的に入力されるようなことができる。

<sup>2</sup> 自然言語文を入力すると、スクリプトを含んだ html のコードを生成することができる。

<sup>3</sup> 2021 年 5 月には、G P T - 3 を使ったプログラム開発のプラットフォームを発表した。自然言語からコードを生成し、用いることができる。

<sup>4</sup> 例えば、バイオケミカルの領域において 6 年間解けなかった結晶構造があっさり解けたという報告もある。

が先導する、ソーシャルメディアやスマホアプリの中では、深層学習を用いた画像認識は当たり前のように使われている。

国内では、このところ、デジタルトランスフォーメーション (Digital Transformation; DX)<sup>5</sup>の議論が盛んで、AI・深層学習の進展についてニュースになることが少なくなっているように感じるが、依然として、深層学習の技術進展は続いている。深層学習は当たり前前に様々なデジタルのシステムの中で使われるようになると同時に、新しい進展が次々と新規のビジネスを可能にしているのである。日本としても、引き続き注視していくべきである。

## 2 我が国の人工知能関係の動き

さて、国内の状況はどうだろうか。この5～6年の動きを振り返ってみたい。私は2015年に、拙著「人工知能は人間を超えるカーディナーニングの先にあるもの」[4]を上梓し、深層学習の重要性を説いた。それもきっかけの一つとなったと思っているが、2015年には、第5期科学技術基本計画に人工知能が重要なテーマとして入ることになった。

その後、人工知能研究を推進するため、いくつかの研究組織が作られた。2015年5月には、産業技術総合研究所のAIセンターが立ち上がり、2016年には理研AIP（革新知能統合研究センター）などが立ち上がった。また、2016年4月には、人工知能技術戦略会議が置かれた<sup>6</sup>。

私は、こういった組織の設立に大きな協力をしてきたと自負している。政治家や役人を前に、何度も深層学習の重要性を説くプレゼ

ンテーションをした。しかし、結果的には、非常に悔しい思いをした。というのは、「深層学習が進展しているから」人工知能に予算を付けるべきであるという予算化されたにもかかわらず、結果的に予算が付いた後の体制は、全く旧来の研究を引きずったものであり、深層学習の研究や社会実装のためのものになっていなかったからである。深層学習という言葉は、枕詞にだけ使われ、実体としての研究や政策の中身は、「人工知能」ならまだしも、「データ」「数理」などの言葉が踊るようになった。これらの違いは、専門外の人から見ると微妙であるが本質的な違いであり、巧妙に意図的にそうした変化が作られ、結果、国や社会は「人工知能」に投資したと思っけていても、結果的には、深層学習ではない旧来からの技術に予算が使われることになった。

何も、これらの内容が重要でないというわけではない。もちろんデータサイエンスは重要なトピックである。数理も重要である。ただ、間違えてはいけないのは、データサイエンスが重要であることは、少なくとも2000年代から分かっていたことである。数理に至っては、それが重要なことはもう何十年前からのことである。つまり、ことさら今、強調する必要がない。国が何か投資するというのは、投資すべきタイミングに来ている対象があるからであって、今回の場合の投資対象は、明らかに急速に進展している（当時、はっきりとその兆候が見えた）深層学習であった。

しかし、人工知能に関する政府のトップ会議には、深層学習の専門家ではない旧来から

<sup>5</sup> もともとDXという言葉は、スウェーデンのウメオ大学のErik Stolterman氏が2004年の論文で使ったのが初めてであるとされている。

<sup>6</sup> 現在のAIに関する動きの状況は、内閣府のAI戦略のページ(<https://www8.cao.go.jp/cstp/ai/index.html>)にまとめられている。

の情報技術の大御所が並び、新設された研究組織も、申し訳程度に深層学習をテーマに含むだけであった。「これは話が違いますよね？」ということ、技術のこともしっかり分かっているはずの文科省の役人に問いただしたとき、あれだけプレゼンテーションをし、深層学習が会議の中心の話題であったにもかかわらず、「深層学習をやるとは書いていない」と開き直られた。ああ、そういうことかと絶望したことは、今でも生々しい記憶である。

政府の人工知能に関する資料を見ていただくによく分かるが、巧妙にキーワードがずらされている。当時、大御所の先生から、「深層学習はただのツールになって終わる」と言われたことはよく覚えているし（結果的には全くの間違いだった）、また、「深層学習が最近傍法（機械学習の最も初歩的な手法の一つ）とほとんど性能が変わらない」と言って回る大家の研究者もいた。これだけ未来のある技術に対して、そんなことを言うのだろうか？と思ったものだ。

なぜ、こういうことが起こるのか？もちろん、日本の各省庁の予算の使い方に関する一般的な問題等もあるだろうが、一つの大きな理由はこの分野の特性に由来する。情報技術の分野は進展が急速であり、深層学習のように新しい技術は若い人しかやっていないからだ。2012年に大きな飛躍を遂げた技術をマスターしている大御所の学者はいない。最も優秀なのは、2012年当時、博士課程に入ったかどうかくらい20代後半くらいの研究者である。それから上は（私も含めて）本当に研究に没頭できる時期にこの技術を習得していない。人間誰しも、自分がよく分からない技術については否定したくなるものである。そして自分が得意な領域に引き込むために、キーワードを「深層学習」から「データ」や「数値」と広げていくのである。

今でこそ、深層学習の意義を否定する人は誰もいないだろうし、今振り返れば、2015年当時、深層学習に投資すべきだったということは誰が考えても分かる。しかし、当時は、深層学習に対して否定的なことを言う人がとても多く、私自身もよく批判された。人工知能研究の未来のため、そして国のためと思っで行った活動が理解されず、様々な困難を伴った。そして、私から見ると、どう考えても深層学習の潜在的な可能性が明らかであるにもかかわらず、技術の真の意味を理解しない人たちによって、学術コミュニティの権力構造によって学術のあるべき姿がネジ曲がっていく姿を見るのは、研究者として大変な苦痛であった。

それから5年経って、結果はどうだっただろうか。多くの若い人が深層学習の研究を始めた。この分野は世界全体で大きく成長し、2015年の時点より、国際会議の論文数も研究者数も何倍にもなった。世界中で顔認証や医用画像に関して、多くの実用化がなされ、巨大なスタートアップがたくさん誕生した。そして、日本は国として数百億円の規模の予算を「人工知能」に投入したはずだが、結果的に、諸外国に大きく<sup>こうじん</sup>後塵を拝している。「人工知能」の論文数だけは数が増えているものの（論文数と研究費は高い相関にある）、日本から深層学習に関する革新的な技術は一向に出しておらず、また、グローバルなビジネスも立ち上がっていない。旧来からの技術を「人工知能」という看板に掛け替えただけなので、革新的な技術が出るはずもない。当然の結果である。

この5年間に日本が被った機会損失は大きなものだろう。恐らく一直線にこの技術に賭けていた場合に比べて、GDP換算で何兆円から何十兆円かは失っているのではないかと思う。恐らくこの分野を<sup>ふかん</sup>俯瞰し、最新の技術

の可能性を誰よりも理解できていたはずの私は、もっとうまくやることはできなかつたらうかと反省の念も多い。

### 3 松尾研究室を中心とする活動

さて、このように、国全体の動きが、本来投資すべき方向と異なる方向に行ってしまった5年間に、私自身が何をしていたかを述べたい。

この5年間に、まず、私は人材育成に力を入れた。革新的な技術の進展期にあつて、まず核となるのは新しい技術を身に着けた人材である。2015年に初めて東京大学に深層学習の講義を作った。これは世界のトップ大学とほとんど遜色ないスピード感であり、少なくとも国内では最も早かった。最初のクラスは30名ほどであったが、それから100名、200名と増え、一気に人気講義となった。これまでに累計で3,000人以上の学生・社会人が受講している。2014年からスタートしたデータサイエンスの講義（グローバル消費インテリジェンス（GCI）という）と合わせ、現在までに合計で7,000人以上の学生・社会人に講義を提供してきた。

2017年には、日本ディープラーニング協会という一般社団法人を作った。この業界が健全に発展し、産業競争力の向上に資することを目的とした団体である。深層学習を事業の核とする企業が集まって構成し、私が理事長を務めている。この分野は前述のように多くの方が「自分は人工知能の技術者である」と名乗りやすく、悪貨が良貨を駆逐する性格の強い領域であること、一方で技術の中心は深層学習であり、それに特化した人材育成が必要であるとの思いから、資格試験を創設した。G検定は、ビジネスパーソンを含め深層学習をビジネスに活用する人材のためのもの、E資格は深層学習のエンジニア向けのものでは

る。これらの試験は、立ち上がり当初から毎年大きな成長を見せ、現在までに6万人を超える方が受験をしている。その合格者の会は約3万人で、人工知能のコミュニティとしては、日本で最も大きなものとなっている。

新しく技術を身につけた若者は、その武器を使おうとする。しかし、大企業も含め、日本のほとんどの組織では、その最新の武器を往々にしてうまく使ってもらえない。自分の専門性と関係ない部署に配属になることも多い。そこで、大企業に就職するのを避け、起業をする学生が増えている。私の研究室では、私がスタンフォードに留学していたこともあり、2008年頃からよくシリコンバレーに学生を連れて行っていた。そうした環境の中で、起業する学生が増え、その中から、2015年には、ニュースアプリを手掛ける株式会社Gunosyが、2017年には、知能化技術の開発を手掛ける株式会社PKSHA technologyが上場した。研究室のメンバーが中心となって、2社、上場した企業が出たことは大きなきっかけになった。多くの方がそれに勇気付けられ、新たな挑戦をするようになった。現在、私の研究室では、起業する、あるいは先輩や友人が作ったスタートアップに参加するということがごく当たり前のことになっており、大企業に行くケースがほとんどなくなっている。

研究室発のスタートアップで、大きく成長を続けているものだけを挙げても、株式会社DeepX、READYFOR株式会社、株式会社ACES、株式会社011o、株式会社ELYZA、bestat株式会社、燈株式会社などがある。研究室発のスタートアップで10社以上、関連するスタートアップを含めると30社近くに上る。最近では、「本郷バレー」の一部として、記事等でもよく取り上げられるようになっている。

日本ディープラーニング協会でも、人材育成からスタートアップを育成するという流れ

を作り出した。高等専門学校生のものづくりの力に注目し、深層学習がものづくりと相性が良いことから、彼らに深層学習の教育を提供するとともに、ものづくりと深層学習を組み合わせたプロジェクトで優劣を競う「高専DCON」を立ち上げた。ベンチャーキャピタリストが付けた企業評価額（バリュエーション）で勝敗を競うというものであり、高専生の技術と世の中のニーズを結び付けようというものである。2019年のプレ大会では全国から18チームが参加し、長岡高専が優勝、その企業評価額（バリュエーション）は4億円だった。2020年、2021年と、参加チームも36チーム、43チームと順調に増え、2021年の大会では、優勝チームの福井高専のバリュエーションは6億円となった。高専DCONはNHK Eテレで毎回放送されている。

そのプロジェクトは、例えば、工場のアナログメーターを自動で読み取る、電線の点検を行う、視覚障害者のための点字の要約を行う、打音検査の判定を自動で行うなど、社会の課題に密着し、市場規模としても十分に大きいものである。そして、彼らの中から実際に起業するものも増えている。長岡高専から株式会社 IntegrAI（インテグライ）、香川高専から株式会社三豊AI開発と Panda 株式会社、そして東京高専からは TAKAO AI 株式会社というスタートアップが生まれた。いずれも順調に成長している。

このように、私の研究室、あるいは日本ディープラーニング協会では、深層学習に関して大規模な人材育成、そして、起業家の育成を行ってきた。そして、強調したいことは、こういった活動は、基本的に研究室の企業との共同研究、あるいは企業からの寄附によって活動を維持しており、国からの経済的な支援をほぼ受けていないということである。

2節で述べた、国からの支援を受けながら、

旧来的な技術に予算を使ってしまう構図と、本節で述べた、国からの支援に頼らず、自助努力で、深層学習という新しい技術と若者に投資をし発展させている構図を比べていただきたい。何も私の周辺の活動を自画自賛したいわけではない。やる気と能力ある若者、そして、新しい技術にきちんと投資することさえできれば、それは社会からも求められることであり経済的にも成り立つし、人材も育て活躍するようになるということである。こうした簡単なことが、いかに日本の国の中でできていないかということである。

#### 4 今後の処方箋

さて、過去は過去として、未来の話をした

い。

一つ目は人材育成である。

いま、全国の大学にAI・データサイエンスの教育が広がっている。「数理・データサイエンス・AI教育プログラム認定制度」が2020年から内閣府で作られている。深層学習の内容も含まれたカリキュラムとなっており、（内容のレベルが高すぎるという問題があるものの）こういった活動自体は大変良いことである。そもそも、こういった教育は、本来あるべき量に比べて圧倒的に不足している。

2012年の深層学習のブレイクスルーから10年近く経ち、良くも悪くも深層学習という技術自体が、新製品や新しいサービスに直結するという時代は過ぎてしまった。以前からあるデータサイエンスや数理とともに、一つの重要な基礎分野として定着しつつある。その意味では、過去5年～10年でいうと深層学習に特化して投資すべきであったものの、今になっては、データサイエンスや数理と併せて教育することにはそれほど違和感はない。

その際に、注意すべき点がある。「若い人が若い人に教えたほうが良い」ということで

ある。今の時代は、情報技術の進展が早すぎて、年輩の教員が新しい技術を教えることが難しい。これを強制すると結局、「看板の掛け替え」が起こる。若い人が教えられるような仕組みにすべきである。極端に言えば、学生が学生に教えても良い。プログラミングからAIまでを教えるフランスの先進的な学校「42」では、学生・卒業生が自治的に学生に教える仕組みを作っている。私の研究室でも、講義を広く提供しているが、学生が講師として学生に教えるというのは日常的である。博士課程の学生が教える場合が多いが、ときには学部の学生が教える場合もある。年齢はもう関係ないのである。

したがって、全国の大学や高等専門学校、高校等にこうした教育の仕組みを広げるときには、若手や学生が教えられるようにすること、動画やオンラインのコンテンツを最大限活用することなどが重要である。

それから、数理・データサイエンスと併せてAIを教えることに違和感はないものの、今の時代に必要な素養をすべてカバーしているわけではない。今の時代には、例えば、ウェブに関する技術は、様々なイノベーションを実現していく上で必須である。それから、日本の様々な組織で変化が必要とされており、「変化を起こせる人材」としての考え方や思考法なども重要である。そして、ビジネスに関しての基礎的な教育も必要である。こうしたことをセットにして提供していく必要がある。

二つ目に、深層学習の研究である。深層学習はこれからも大きく進展する。この可能性を過小に見積もってはいけない。現在の深層学習は、ややタコツボ化しつつあり、局所的

な小さな改善が多くなっている。また、一部の領域では明らかな限界も見せ始めている<sup>7</sup>。しかし、深層学習の根本的なアーキテクチャからの変更が今後起こるはずであり、そうした先には、また急速な進展が起きると思われる。

私の研究室では、2021年に「世界モデル・シミュレータ寄付講座」という寄付講座を立ち上げた。実空間の把握、そして、そこでの相互作用に関して、技術が大きく進展し、それが機械やロボットへの応用の基盤になると同時に、自然言語理解の基盤になるはずというのがその目論見である。そして、脳科学と人工知能が融合する領域でも大きなブレークスルーが起こる。物理や化学など、基礎的な学問領域における深層学習の活用についても、ブレークスルーが起こる。したがって、こういった分野への投資は継続していくべきである。

科学技術の研究に対する国の投資の一般的な問題点として、社会のニーズや事業について考えるのが苦手な研究者に、応用について書かせるべきではないし、それに基づいて審査をすべきでもない。あくまでも、研究者は知的好奇心と自らの専門性に基づいて研究をすべきであり、国はそれを広く薄く支援すべきである。一方で、今回の人工知能・深層学習のように、国がその技術を一気に進展させたいときは、むしろ研究ではなく、開発、社会実装を中心とした支援をすべきである。その際に、ビジネスの感覚に優れた経営者や投資家等に、技術のことを理解してもらった上で、その方向性を決めていくべきである。つまり、長期的で本質的な基礎研究と、それを社会に活用するための仕組みを大きく二つに分けるべきであり、それをまぜこぜにして研

<sup>7</sup> 例えば、OpenAIは、ロボット研究からの撤退を2021年に発表した。

究者にやらせるべきではない。これは、科学技術予算全般に関しての問題提起でもある。

三つ目に、産業の育成である。基本的に、今の深層学習はどんどん使いやすくなっており、プログラミングができる人であれば簡単に使うことができる。企業の中のみならず、経営者レベルでも、AIのプログラミングを少しでも試したことがある人の数を増やしていく必要がある。ものづくりで、はんだごてをまず触ったことがないといけないように、今の深層学習を使うには、少しでもプログラミングをしなければならない。今の議論を見ても、そういった体験をすることなしに、AIやDXを語っている人が多すぎるように思う。日本のものづくりでは、昔から「わいがや」で意見をぶつけ合いながら技術が進んできた。深層学習・人工知能も、多くの人が実際に動かすようになると、同じような雰囲気になる。

日本の様々な地方で深層学習、人工知能そしてより広くデジタルの技術を活用することができる。そもそも、ものづくりと深層学習は相性が良い。農業や林業、水産業、観光業などとも、深層学習は相性が良い。現場があって、その情報を把握しないとイケないときに、画像の情報が大いに役立つからである。だからこそ、高専DCONを作ったわけである。したがって、日本の様々な地方で、人工知能の教育をし、それを実践的に活かすことを学び、地方の産業活性化に役立てていくことができるはずである。

そのためには、きちんと事業として顧客に付加価値を出して「儲ける」ということが第一である。そして、そのために必要な人材の要件が決まる。そのために必要な教育が決まるという順である。日本では、出口を意識し

ない教育の議論が多すぎるように思う。

スタートアップの育成を、日本全体のDXの取組とも合わせて、取り組んでいく必要がある。私の研究室では、2021年から「AI経営」という寄付講座をPwCの支援の下、立ち上げた。AIを経営に活かす方法は、世界中で事例がたまっている。人事、会計、法務などに活かす方法、医療や製造、物流などの分野で活かす方法など、水平・垂直双方に様々な取組がある。これを、ハーバード大学やMITなどのトップ大学では、講義として提供し始めている。これと同じように、AIを使ったビジネスのケーススタディを学生に提供するというのがこの講義の趣旨である。第1回目の講義が始まっているが数百人の学生の登録があり、人気を集めている。

また、スタートアップを、より大規模に輩出していくため、2021年からは「起業クエスト」というものを研究室で立ち上げた。これは、今までの起業の成功パターンを仕組み化したものであり、AI・データサイエンスの教育（ステージ1）から、企業との実践的な共同研究・開発のプロジェクトを経験し（ステージ2）、起業のための知識を得て（ステージ3）、実際に起業するという、一連の流れをプログラム化したものだ。これによって、より多くの起業に挑戦する人が生まれ、またその成功確率が上がってくることを目指している。

私の研究室の活動だけでなく、東京大学でのアントレプレナーシップ教育は盛んである。産学協創推進本部を中心として、長年講義を提供しているアントレプレナー道場、駒場生向けのアントレプレナーシップ・ブートキャンプや新入生向けのアントレプレナーシップセミナー、インキュベーションの施設である本郷テックガレージやFoundX、そしてEDGE NEXTやToday To Texasといった特徴あ

るプログラムなど、幅広い。私も、東京大学工学系研究科の講義として、「データ駆動型起業演習」「データ駆動型事業立案演習」などの講義も行っている。2021年には、アントレプレナーシップ教育デザイン寄付講座も立ち上がり、起業家の育成がますます盛んになっている。こうした取組を、全国の大学等にも広げていくことが必要だ。

以上、人材育成、研究、そして産業の育成について述べた。いずれにも共通するのが、若者と新しい技術、この二つをしっかりと信頼し、期待することである。そうすれば、世界的には「普通の」イノベーションが「普通に」

起こるはずである。日本だけが30年間も負け続けるというようなことは、普通にやっていたら起きないはずである。地方からもグローバルな企業が、一定の確率で新たに誕生してもおかしくない。旧来の技術にこだわるとか、若者の芽を摘んでしまうといった阻害要因をなくす。そして、若者と新しい技術を信頼し、期待して待てば良い。それがやるべきことである。

現在の日本にもうそれほど余裕はないはずである。当たり前のことを当たり前にしていけば、必ず成長が生み出せるはずである。

#### 【参考文献】

- [1] Vaswani, A. (2017). Attention is all you need. Proceedings Annual Conference on Neural Information Processing Systems (NIPS 2017), 6000–6010.
- [2] Brown, T. et al. (2020). Language models are few-shot learners. arXiv:2005.14165 <https://arxiv.org/abs/2005.14165>
- [3] Jumper, J. et al.: Highly accurate protein structure prediction with AlphaFold., Nature 596, pp. 583–589, 2021
- [4] 松尾 豊『人工知能は人間を超えるか—ディープラーニングの先にあるもの』(角川E P U B選書) KADOKAWA (2015)