



**JAPANESE
CIVILIZATION
INSTITUTE**

日本文明研究所

24-5 Sakuragaoka-cho, Shibuya-ku, Tokyo 150-0031
Tel: 03-5456-8082 Fax: 03-5456-8388
Mail: info@japancivilization.org
<http://www.japancivilization.org/>

ニューズレター No. **12** 2018年夏号

2018年5月23日、日本文明研究所の第12回シンポジウムとして、特別講演「人工知能は人間を超えるか!？」が行われました。登壇者は、人工知能研究の第一人者である、東京大学准教授の松尾豊氏。モデレーターを作家で当研究所所長の猪瀬直樹氏が務めました。深層学習（ディープラーニング）の急激な進化により、世界が劇的に変わろうとしているいま、知っておくべきこと、日本の人工知能研究の現状や、日本の産業がどのように生き延びていけるのかなどが熱く語られました。その模様を一部載録いたします。

『週間読書人』2018年8月10号 第3251号ならびに『WEB読書人』より転載)



ごあいさつ

日本文明研究所理事、株式会社紀伊國屋書店代表取締役会長兼社長 高井昌史

当研究所の創設以来、理事を仰せつかっておりまして、本来ならば、都築仁子理事長がご挨拶すべきところ、代理を務めさせていただきます。本日はAI研究の第一人者である松尾豊先生のご講演ということで大変楽しみにしております。松尾先生は様々な著作を上梓されており、研究書のみならず、『マンガでわかる！人工知能』『よくわかる人工知能』といった一般読者向けの平易な解説書もお出しであり、いずれも売れ行き好調です。本日のテーマは「人工知能は人間を超えるか」です。個人的には、社長職をAIに委ねたら経営は必ず良くなるのだろうか？というところを、ぜひ伺いたいと思っております。せっかくの機会ですから、皆様どうぞ積極的に先生にご質問をお寄せ下さい。

AI時代に求められる「霊性」について

会長 後藤俊夫

今回のテーマであるAIと日本文明の関連について考える中で、浮かんできたのが、100年も前、禅について英語で世界に発信した鈴木大拙のことです。AIは人間の理性、知性、いわゆる左脳の部分を超えるとも言われます。知識、論理、推論まで含めて、創造性の領域にも接近し、その先の感性部分に到達するべく、日夜、研究が進んでいます。数年前、「日経星新一賞」(日本経済新聞社主催)という文学賞の候補作にAIが書いた作品が残ったことが話題になりました。

すなわち、文学の世界でも、人工知能が人間と同等、あるいはそれ以上のレベルに達する可能性がある。音楽も同様で、ネット上にはロボットがつくった音楽が多数流れています。われわれの脳内で左脳が司る理性、右脳が司る感性の両方がAIに追い抜かれかねない。その上で最後まで残る分野はあるだろうか。そう考えて浮かんだのが、鈴木大拙の「霊性」という言葉です。

霊性とは何か。鈴木大拙は、われわれの精神の真ん中にある、包みきれない何かと、そこに付随する宗教意識のような

部分を霊性と呼びました。個別宗教については言及せず、絶対者の存在という意味での宗教的存在です。それについて、「精神が物質と対立して、かえってその桎梏に悩むとき、みずからの霊性に触着する時節があると、対立相克の悶えは自然に融消し去るのである」

と述べています。人間が人間として悩まないとき、すなわち、赤ん坊のときには、霊性というのは目覚めない、と。加えて、日本という国が時代として、霊性に目覚めたのは鎌倉時代であり、その中の一つが禅だとも書いています。

「鈴木大拙」と聞くと禅の専門家と考えがちですが、よく調べると生涯をかけて霊性、特に「日本的霊性」の研究に没頭したというのが正確なところのようです。AIが理性、感性の域に近づきつつあるという今、われわれがやるべきことは、「人間」をもっと知ることではないでしょうか。AIの時代になればなるほど、文明というものの重要性も問われるのではないのでしょうか。そして、知識や感性のさらに奥底、鈴木大拙のいう霊性の領域にある魂のところが問われるのではないかと感じております。「日本的霊性」ということを頭の片隅に置きながら、松尾先生のお話を伺うとこれからの一つの手がかりになるのではないかと考えております。

特別講演

人工知能は 人間を超えるか!?

モデレーター

松尾豊

×

猪瀬直樹

ディープラーニングの 大きなイノベーション

松尾 二年前、Google Deep Mindが開発した人工知能のアルファ碁が、韓国のプロ棋士・李世石^{イセドク}九段を破りました。また昨年は中国の柯潔^{カキョ}九段に勝ち、太刀打ちできる人間は誰もいない状況です。アルファ碁はプロ棋士の棋譜データを学習し、それをベースに自己対局を繰り返すことで強くなっていたのですが、さらに、昨年十二月に開発されたアルファZeroは、学習用の棋譜データすらいりません。そして囲碁以外のチェスでも将棋でも、自己対局で強くなることのできるのです。将棋では最強の人工知能ポナンザが、昨年名人に勝ちました。が、さらに世界コンピュータ将棋選手権で、ポナンザにelo(エルモ)が勝利し、そのeloのレベルにさえ、アルファZeroは、真つさらな状態から学習をスタートして、わずか二時間で到達してしまいます。名人が一生かかって身に着けるものを、わずか二時間で到達する。こうした人工知能に中心的に使われている技術が、ディ

ープラーニングなのです。

いま人工知能が大きなブームですが、実はこれは三度目のブームに当たります。人工知能の研究は一九五六年に始まり、六二年を経ています。第一次AIブームから「人工知能は人間を超える」といった妄想が語られてきました。ところが技術が追いつかず、妄想が失望に変わり、冬の時代がくる、ということは何度も繰り返しています。今回のブームでは、人工知能は何ができて何ができないのかを、きちんと知る必要があると思います。

実はいま騒がれている自動運転技術は、十数年前から研究されています。大まかに言えば、いまある技術は昔からのものが少しずつ良くなつて今に至ります。そして、昔出来なかったことはほとんどの場合、いまもできないままなのです。でもその中で、ディープラーニングの技術だけは、大きなイノベーションが起っています。

では、ディープラーニングでは何ができるのでしょうか。一つ目は「認識」です。次に「運動の習熟」。それから「言葉の理解」、この三段階です。



松尾豊氏

「認識」から話していきますと、猫、犬、狼、三枚の写真を、人はパッと見わけることができますね。ところがこれが、コンピュータには難しい。まず「目が丸ければ猫」だと、分類の定義を与えることにします。次に「眼が細長い場合に、耳が垂れていれば犬、耳がピンと立っていれば狼」と判断基準を与えます。一見、このルールで認識できそうに思いますが、世の中にはいろいろな犬がいて、例えばシベリアンハスキーは、耳が尖っているけれど狼ではない。犬っぽさ、狼っぽさを定義するのは、

なかなか難しいことなのです。

こうした「うっぽさ」を「特徴量」と言いますが、特徴量を人間が定義している限りは、画像認識の精度は上がりませんでした。人間は、「うっぽさ」を、無意識のうちに学習しています。コンピュータも、人間のように特徴量を学習できるシステムが必要だった、ということなのです。それを可能にしたのがディープラーニングです。「Googleの猫」が有名ですが、YouTubeから取ってきた画像をコンピュータに大量に見せるだけで、猫らしさが学習でき、そういう仕組みが開発されたのです。

二〇一二年は、画像認識のエポックメイキングでした。世界中で同じデータセットを使い、数万枚の画像をコンピュータが当てて、その正答率を競います。が、なんと、二〇一二年の優勝チームは、従来の記録をいきなり十%も下回る、十六%のエラー率を叩き出したました。それは、それまで人力で定義していた「特徴量」を、ディープラーニングが自動的に学習するシステムが開発されたためでした。これ以降、画像認識のエラー率は、速度を上げて低下して



猪瀬直樹所長

いきます。二〇一三年には十一%、一四年は六・七%、現在は二・三%です。同じタスクを人間が行うと、五・一%のエラー率だと言います。既に画像認識においては、コンピュータが人間を超えているということなのです。こうした変化が、わずか五年程の間に起こっているのです。

画像認識には、画像中の物体を取り出す物体検出や、対象物の領域を切り出すセグメンテーションなど、さまざまなタスクに細分化されます。例えば、YOLOv2という、物体検出プログラムでは、一つの映像の中

に映っている個々の物体それぞれを、一秒間に四七フレーム読み取ることができます。しかもこのプログラムは全世界に共有されており、ネット検索ですぐに手に入れることができます。

認識から運動の習熟、言葉の意味理解へ

この画像認識をロボット技術と組み合わせると、「運動の習熟」が可能になります。人間は同じ行動を繰り返すと、うまくできた行動を強化することで、だんだん上達していきます。「強化学習」と言いますが、それと同じことを、コンピュータにやらせる仕組が開発されています。例えば、ブロックの凸と凹を組みあわせたり、蓋を開け閉めしたり、様々な高さの棒にハンガーをかけるなど、いろいろなタスクを、目で認識しながら調整していきます。十数台のコンピュータがカメラで運動の様子を撮影し、経験を共有することで、高速に自ら学習します。従来のロボットや機械の技術では、形やサイズが違うものを、調整しながら持

つということはできませんでした。画像認識の精度が悪くて眼が見えなかったからです。

物によって持つべき場所が違うことを、人間は、赤ちゃんのときに、いろいろなものを持つては落すことを繰り返し、いつの間にか学習します。それと同じことを、ロボットができるようになってきています。

「モラベックのパラドックス」と言いますが、コンピュータにとつては子どものできることほど難しいのです。コンピュータによる医療診断やチェスゲーム、数学の定理の証明などの研究は、そこそこできるようになっていましたが、三歳児でもできるようなこと、例えば画像認識はなかなか精度が上がらなかった。それは人が、自分がそれをどうしてやるのか、説明できないからです。

第三次のAIブームによって、この状況が変わりました。つまり最近のAIは、高度なことができてすごいのではなく、三歳児でもできるような簡単なことが、できるようになってきたからすごいのです。それによって、世の中がどう変わっていくのかを、考えていく必要があります。

認識、運動ときて、次は「言葉」

です。これは子どもの発達の過程とよく似ていて、見て分かるようになると、体の動かし方が上手になり、いろいろなものを触ったり、落したりし、概念を掴めるようになる。そして言葉の意味が分かるようになります。

僕は二〇一四年の秋の段階で次のように予測しています。人工知能技術は①画像認識の精度向上 ②行動予測 ③環境変化に影響しにくい行動の学習 ④周りの環境に上手にインタラクションしながら行動する技術 ⑤言語理解 ⑥大規模知識理解という、六つのステップで発展し、二〇三〇年ぐらいまでに完成する。①②で認識ができるようになると医療画像の診断、防犯・監視、③④の運動の習熟までいくと自動運転、物流・建設の監視や、農業の自動化、あるいは介護、調理・掃除、それから⑤⑥の言葉の意味理解までいくと、翻訳や、ホワイトカラー支援全般ができるようになるのではないかと。実際に、その通りに進んでいます。ただ予想を外したのは、三〜五倍早いスピードで進んでいるということです。

コンピュータは現段階で、言葉の

意味理解もできるようになりつつあります。一つはイメージキャプションという、画像を見て文を生成する技術です。例えば、男性がギターを弾いている写真を取り込むと“man in black shirt is playing guitar”という文が出てきます。逆に、文を入れると絵が出てくるという技術もできつつあり、“A very large commercial plane flying in blue skies.”という文を入れると、大きな飛行機が、青い空を飛んでいる画像が出てくるのです。

ここでは、画像を検索するのではなく、文書を元にイメージが作り出されています。私たちは何かを聞いたとき、頭の中に無意識にイメージを描いています。それと同じことが行われているのです。

Google翻訳は、二年前からディープラーニング方式に変わり、非常に精度が上がっています。例えば英語からスペイン語の翻訳や、フランス語から英語への翻訳は、人による翻訳とほとんど遜色ないところまで到達しています。英語から中国語や日本語となると、言語の体系がだいぶ違うので、人間にとっても難しいですし、人と自動翻訳との差も

出ています。が、それでも従来より人間の能力に近づいています。

ただもちろん、いかに精度が上がっても、戸田奈津子さんに替わるわけではありません。人は、心の動きや文化的な背景も含めて、翻訳しているからです。

同様に、この先自動機械が、人の心を打つ小説や俳句作品を作れるかといえば、それは難しいでしょう。小説や俳句は、感情や本能が表現されたものであり、それらは生命由来の現象です。人間と知能の関係は、鳥と飛ぶこととの関係だと思ってください。工学的に飛ぶ原理を解明して飛行機が作られますが、飛行機は朝になつてもチュンチュン鳴かない。それと同じで、人工知能は生命的な感情や葛藤は持ちません。人間の知能を工学的に解明したものが、人工知能なのです。

「眼の誕生」による、 人工知能の進化

この発展の理由を一言でいうと、「眼の誕生」です。

古生物学者アンドリュウ・パーカ



「の『眼の誕生』という、カンブリア爆発について書いた本があります。カンブリア期とは、地球四六億年の歴史の中の、わずか一千万年の間に、現存する生物の全ての種が出そろった時期だと言われます。なぜカンブリア爆発が起こったのかは諸説あるのですが、パーカーは、眼が出来たからだ」と、「光スイッチ説」をとえています。それまでの生物は眼がなかったので、嗅覚や触覚を頼りに、のろろ進んでぶつかると食べたり、ぶつかられると逃げたり、緩慢にしか動けなかった。ところが、史上初めて高度な眼を持つ三葉虫が現われて、生存上、圧倒的に有利なので大繁殖した。そのうち逃げる方も眼を持ち始め、早く泳いで逃げようとか、みつからないように擬態しようとか、眼を持ったことにより、生物の生存戦略が多様化し、種が多様化したと僕はこれと同じことが、ロボットの世界でも起こると考えます。眼を持つようになったことで、出来るタスクが圧倒的に増え、機械の種類も急速に増えると。

カメラのイメージセンサーは、生物でいう網膜に当たります。我々は網膜で受け取った信号を、脳の視覚野で処理することによってものを見ています。この視覚野が、ディープラーニングにあたります。今までの我々の文明は、眼の見えない機械をいかに使いこなすか、その工夫の歴史でした。工場で扱われる品物は、形も材質も定まっている必要があったし、工場のラインも、眼の見えない機械に合わせて管理されていた。そして現在自動化されていないのは、眼の見えない機械では作業が難しい分野です。その典型が、農業や建設、食品加工業などです。

トラクターもコンバインも、動かしているのは人間です。稲やジャガイモのような、根こそぎ取るものは、まだいいですが、みかんやりんご、トマト、キュウリなどの果樹や園芸作物は、人間が眼でみて収穫する必要があります。農業は、ほとんどの作業に眼を必要としています。だから人手不足で困っている。建設現場も、鉄筋を組んだり、溶接をしたり、コンクリートを流し込んだり、全て眼が必要です。食品加工も、眼がなければ、素材を切ることも、フライパンにのせて色が変わったらひっくり返すこともできない。

しかし眼の技術によって、農業や

建設、飲食産業も、自動化できる可能性が出てきました。これは巨大なインパクトです。例えば農業の中では、トマト収穫は比較的自動化しやすいし、コストの見合う分野ではないかと思えます。建設業の中では、溶接は比較的単一作業でし、溶接工が足りないので、実現可能性が高いでしょう。調理はかなり複合的な作業で、最初から自動化するのは難しいですが、食洗機に皿を入れるぐらいは出来ると思います。そのように可能なところから自動化し、周辺領域へと広がっていくのがよいのではないかと思えます。

実は、農業機械、建設機械、食品加工機械に日本は強く、世界にもシェアを広げています。ハードウェアとディープラーニングの技術をうまく擦り合わせて、自動化機械を開発する。これは日本企業が力を発揮しやすい分野だと思います。日本はインターネット関連のビッグビジネスを作れずに、ここ二〇年経済的に苦しんできましたが、眼を持つ自動機械の開発に関しては、うまくすればいいポジションに付けるのではないかと。

第一次産業革命では蒸気機関がで

き、それまで人間あるいは家畜の筋肉に紐づいていた「力」が、切り離され、代替されることとなりました。同じように、機械が眼を持つということは、「認識」が人間から切り離されて、社会の必要なところに再配置されるということです。

実際に一番進んでいるのは医療分野で、X線、CT、MRI、それから皮膚科、内視鏡などの診断は、軒並み医者の精度を超える自動機器が開発されています。つい先月、アメリカ政府機関のFDAが、ディープラーニングによる眼底検査機器を認可しました。もはやサポートではなく、医者の業務を替わって行うところまでできているのです。

次に進んでいるのが顔認証です。iPhone Xにも入っていますが、中国のテンセントの入退出の管理は全て顔認証です。中国の全国土には一億台以上の監視カメラがあります。ほとんどディープラーニング化しています。モニター室で人が見ていする必要はないし、尋ね人が最後にどこかのカメラに映ったかを十秒で探し出すのです。中国は、国民を監視したいですから、非常に相性がいいのですが、こうした状況も、今

後どう進展していくのか注目されると思います。あるいは介護施設や病院で、二四時間、機械の眼が見守ることもできますね。

現在は、お掃除ロボットはありますが、片づけロボットはありません。片づけとは物を認識して、元あった場所に戻す作業だからです。でもこれもいまの技術ならば、製品化の可能性が充分あります。片づけロボットが日常化すれば、朝出かけて夜帰ってくる、全ての物が元の場所に戻っている、「家のホテル化」が起こります。生活感は大きく向上するでしょう。

ディープラーニングは、最小二乗法である

ディープラーニングとは何かというのを、構造からもう少し説明します。ディープラーニングとは一言でいうと、「深い関数を使った最小二乗法」です。最小二乗法は統計学で使われる方法ですが、例えばエクセルで散布図を描き、近似直線を追加すると線が引けます。これを引くアルゴリズムが最小二乗法です。

例えばある店の、その日の気温と飲料の売り上げとの関係を散布図に書きます。近似直線の追加で線を引き、 $\langle y, \|a + bx \rangle$ という式になります。この y は推測値を表し、実際の値とどれだけずれているか、つまり各点がこの線からどれだけ離れているのか、誤差を取ります。プラス方向とマイナス方向の誤差と両方あるので、誤差を二乗して足し合わせます。二乗和は全ての点で、線から多少ずれているので、これを最小にするように a や b の値を決めると、この点の上にうまく乗るような線がみつかるのです。これが最小二乗法です。

今は気温という一変数に対する、飲料の売上で計算式を作りましたが、これを気温と湿度の二変数にします。すると三次元になり、線が平面になります。そうなってもやることは一緒で、 y を X_1 と X_2 という二変数を使った式 $\langle y, \|a + bX_1 + cX_2 \rangle$ で、実際の値とどれだけずれているか、誤差を二乗して足し合わせ、この誤差の二乗和が最小になるように、 a や b や c の値を決めていきます。いま一変数を二変数にしましたが、これを一変数にしても同じです。

1万変数の問題というのは普通はなさそうですが、例えば 100×100 の画像に猫が映っているかどうかを判定したいは、この問題にあたります。つまり一万個の変数に対し、猫なら1、猫でなければ0を返すような関数を見つけるようなことがディープラーニングでは行われます。そのときにもやることは同じで、 $X1$ から $X10000$ までの一万個の変数を使った式 $\langle y \rangle = K0 + K1X1 + K2X2 + \dots + K10000X10000 \rangle$ を作り、実際の値との誤差の二乗和をとり、それを最小にする $K0$ から $K10000$ までの値を決めればいい。このように、猫の画像認識は、最小二乗法で学習できるということです。

ディープラーニングは、経済学やマーケティングの、重回帰分析や多変量解析に近いものです。ただ普通、重回帰分析とか、多変量解析は、多くても三十〜四十変数ですが、ディープラーニングでは、数万変数から数億変数使います。データがものすごくたくさんありますし、パラメータ（変数値）を推定するために、計算機のパワーがものすごく必要です。つまりディープラーニングとは、最小二乗法のお化けだと思ってください

い。

さらに言うと、ここまでの説明はやや簡略化したために、少し事実と違うところがあります。ディープラーニングで活用するのは、線形の一次式ではなく、深い式です。 $X1$ や $X10000$ などの変数から、直接、関数を定義するのではなく、一度別の中間的な関数 $z1$ や $z2$... $z100$ などを定義して、この関数を用いて、猫関数を作るのです。式は $\langle \text{猫}(X) \rangle = K0 + K1X1 + K2X2 + \dots + K10000X10000 \rangle$ となり、中間的な関数を増やすほど、二層、三層、四層、五層...と層が増えることになります。でもやることはやはり同じで、猫関数を非常に深く定義したパラメータは、最小二乗法でみつけることができるのです。

一次式にいくら中間的な関数を定義しても線形のままですが、ここに非線形な要素を入れることで、刺激が一定量を超えると発火するという神経細胞ニューロンの働きを模擬しようとしています。Xの値が正のときには1をとり、負のときには0をとるシグモイド関数 σ を作り、この σ をそれぞれの関数の頭につけることで $\langle f1(X) \rangle = \sigma(K01 + K1X1 + K2X2$

$+ \dots + K10000X10000) \rangle$ 、中間的なそれぞれの関数が非線形性の強いものとなり、全体の猫関数も、非線形性の強い関数となるのです。ただこのパラメータも基本的に、最小二乗法で求めているということです。

このように関数を深くするのは、我々の住む世界が階層的だから、関数も階層的な方がいいということだと思います。また、ありとあらゆる画像が構成する空間を考えると、そのなかで「猫」が該当する領域は部分空間になります。その領域はぐにやぐにやとねじまがった複雑な形をしています。こういう形を切り取るために、表現力の高い関数を作る必要がある。表現力が高くなるには、通常は、パラメータの数が多くなりますが、そうするとデータがたくさん必要になりますので、できるだけパラメータを少なくしながら表現力を上げることを考えたときに、最も効果的な方法が、簡単な関数を重ね合わせることなのです。

深い階層を持った関数を作って、パラメータからデータを推定することとは、何十年も前から研究者が目指していたことですが、なかなか実現しなかった。その理由が、計算機の

パワー不足とデータ不足。それがいま解消され、お化けのように大量のパラメータを持つ最小二乗法が解けるようになった、ということです。

美しき シグモイド関数から、 現実的なReLUへ

現在はさらに、シグモイド関数からReLU (Rectified Linear Unit) という活性化関数を使うようになってきています。ReLUの式は $\langle \max(0, x) \rangle$ で表し、負の領域のときは0で、正の領域になるとXに変わります。なぜシグモイド関数よりReLUがいいかというと、シグモイド関数は、入力値Xが大きくなればなるほど、出力値Yが1に近づき張り付いて動かなくなるのです。微分して最適値を求めようとしますが、ほとんど微分が0になってしまい、最適解に近づけないのです。しかしReLUの値はXですから、微分したときに必ず定数項が残るので、うまく最適解に収束します。

ReLUが使われるようになったのは、二〇一二年頃からで、人工知能



のニューラルネットワークの研究が始まってから何十年もの間、研究者はシグモイド関数ばかり使っていました。なぜなら、シグモイド関数の $\sigma(x) = 1/(1+e^{-x})$ という式が美しいからです。ところがReLUの $\max(0, x)$ は格好がよくない。ディープラーニングは、以前は現実的ではなかったのですが、どうせできないならば、理論的に美しい方を使いたいと考えたわけですね。

もう一つだけバッチ正規化 (Batch Normalization) を紹介します。ニューラルネットワークで深い関数を最適化するときに、どのぐらい間違えたのか誤差を戻していく、誤差逆伝播という方法があります。変化の割合が0に収束してしまったり、無限大に発散したり、これまで深い関数はなかなか最適化できないとされてきたのですが、バッチ正規化では、各ニューロンが出力する値を平均0、分散1になるように、調整しましょう。要するに各ニューロンが出す値を、偏差値にしよう。それだけのことで、それでうまくいくようになったのです。微分が0に収束したり、無限大に発散することがなくなり、どのニューロンもそれなり

の偏差値、つまりある一定の値を返すと。非常に最適化の性質が向上したということです。

これが提案されたのは、二〇一五年のことでした。研究を何十年も続けていて、誰も正規化しようとしなかった。これは統計に携わる人なら、普通に誰でもやるようなことなのです。皆できないかと思うんですけど何十年も停滞してきましたが、その壁を越えたら、どんどん発展が進み、現在のような精度の向上につながっているということです。

ディープラーニングは、日本人向きの技術？

ディープラーニングは、一言でいうと、Xの値を入れるとYの値が推測できる「最小二乗法」です。それを変数一万から一億使い、簡単な中間関数を重ねることで非線形なディープな式とし、神経細胞ニューロンの働きを模擬して深層学習を可能にします。

ともあれ最小二乗法ですから、XとYに何をもってくるかというだけの話なんです。Xに画像を持てき

て、Yにクラス——猫や犬をもつてくると、画像認識ができる。またXに画像を、Yにクラスとバウンディングボックス（画像を囲む長方形の枠線）を出力するように設定すると、

物体検出のアルゴリズムになりますし、Xに英語の文をもつてきて、Yに日本語の文をもつてくれば、翻訳になる。それをどのように、ニューラルネットワークでつなぐかということを考えれば、基本的に何でも学習できる。原理としては簡単です。

僕は、原理が簡単で、細かいチューニングがたくさんあるディープラーニングの設計は、日本人向きだと思っています。回路を細やかに修正したり、コッコッパラメータのチューニングをしたり、データを集めれば集めるほど、精度が上がる。努力が報われる作業なのです。それから、ディープラーニングの設計とハードウェアの設計の、組み合わせ方によって大きな価値が生まれますから、その辺りも日本人の得意とするところだと思います。

ところが現状としては、ディープラーニングの技術において、日本は全く世界に追従できていません。ディープラーニング関連の学術論文一

五〇〇本を引用回数順に並べると、一番トップは一七〇〇回引用されています。画像認識に関する二〇一二年のジェフリー・ヒントン教授たちの伝説的な論文です。

さらに、これを研究者別の引用数ランキングにすると上位に、ヒントン、ヨシユア・ベンジオ、ヤン・ルカンというディープラーニング界の教祖たちが並びます。彼らは冬の時代はずっと研究し続けて壁を突破した、別格の人たちです。がそれ以外は、二十代後半から三十代前半の若い世代ばかりです。しかもほぼ二〇一二年以降の論文です。他の学術分野で、論文の引用数順で並べ替えたときに、上位にこれだけ若い人が入ることはないですよ。ディープラーニングがいかに従来の技術と隔絶しているか、全く新しい分野が出てきたことを示していると思います。

そして、このランキングに日本人は全く入っていません。一番上が四〇〇位です。残念なことに日本では、国も企業も、ディープラーニング以前の人工知能にばかり投資しています。そしてこの勝負は人材獲得で決まるのです。ディープラーニングの優秀な人材はほとんど皆、若い。先

ほどのリストで三〇〇位前後でも、初任給の相場が年収五千万円以上です。これが三十位以内になると、桁が一つ上がり、中には数十億円の年俸をもらっている人もいます。サッカーのトップ選手のような扱いです。Deep Mindには、こういうトップレベルの人材が百人以上いますし、GoogleやFacebookには、数百人単位でいるのです。そして、中国のアリババやテンセントは、いま急速に人材を集めています。その闘いに日本は全く参加していないという状況です。

日本の問題点はまず、古いこと。ディープラーニングの急速な進展についていけない。インターネットが現われたときと同じように、若い人でディープラーニングに取り組みようとしている人はたくさんいます。その価値を、人事権や運営権を持つ年配者が理解しないと、昔ながらの分野への投資になってしまふ。

二つ目に、遅い。製造業にとつて大きなチャンスなのに、日本企業では軒並み意思決定ができない。僕は「消極的全張り」と呼んでいます。意思決定しないがために、全てに薄く張る。大手企業に人工知能を担当

する部署があつても、あるだけで意味がない。

三つ目が一番重要ですが、人への投資になっていない。国が予算をつけても、最後はスーパーコンピュータを買ってしまう。というのも、今の日本の人事制度では、若い人に数千万を払うことができないのです。

僕は、東大で二〇一五年から、ディープラーニングのゼミを始めて、初年度は七〇名が受講、翌年は一九〇名、翌々年は三六〇名と倍々で増えています。二〇一五年は、グローバルなトップ大学が、ディープラーニングの講義を始めたのと同様なので、その点では遜色ないと自負しています。技術を身に着ければ明らかなチャンスがあるので、起業する人がたくさん出て、松尾研究室から今まで二社が上場しています。

産業的な仕組み作りと人材育成

これは非常に大きな技術革新なので、産業的な仕組みが必要だと思い、昨年六月に日本ディープラーニング協会を設立しました。その目的の一



猪瀬 直樹

つは、ディープラーニングの技術者を増やすということです。松尾研のみで取り組んでも高が知れていますから、他大学や様々な事業者に、学生から社会人まで、幅広く学ぶ機会を提供して欲しいと。

そして、人工知能は偽物と本物の区別がつきにくい領域なので、資格試験を用意しました。一つはディープラーニングの技術のある一定以上身に着けた人に出す、エンジニア資格（E資格）です。もう一つはディープラーニングの基礎的な理解を、ビジネスに生かす力をはかるG検定です。

日本のIT業界が産業競争力にながらなかつた大きな原因の一つは、発注者である国や地方自治体、医療機関、金融機関など、ユーザー側のディープラーニングのリテラシーの低さにあると思っています。そのため、単に工数による金額設定が成り立ってしまった。これがレストランならば、一万円する料理だからうまい、と言っているようなものです。普通は顧客が、安いけれどもうまい店を見つければいい中で、店側はしのぎを削って業界の質が上がるわけですよ。

ディープラーニングで、産業を変えていこうと思うのなら、発注者側のレベルを上げる必要があります。そういう発想からG検定を作りました。これは、人工知能を使った事業を考えている経営者、新規事業、人事、マーケティング等の担当者に、ぜひ使っていただきたい。一、二週間勉強すれば、取得できる資格です。そして取得すればITベンダーと話すときに、重視すべき点がかかるようになります。第一回のG検定は昨年十二月に実施し、受験した約一五〇〇名のうち半分が合格しました。今後、年三回実施予定です。

ディープラーニングを用いたものづくりは、日本にとって大きなチャンスだと思います。少子高齢化に起因する労働力の不足は、「眼」の技術で解決できる確率が高まっています。しかしこのスピードのままだと、アメリカと中国が闘いに勝ち、日本は出る幕がないでしょう。それをここから巻き返そうとするときに、いかにディープラーニングに投資するか、それをいかにスピード感をもってやるか、真剣に考えていかなければなりません。

Dialogue 日本に、 ディープラーニング 総力戦研究所を！

猪瀬 ディープラーニングの技術を高めることは、日本人に向いている。しかし日本のシステムは、年功序列、終身雇用で、五〇〜六〇代に決定権がある。世界のディープラーニングの先端をいく技術者は、ほとんど二〇〜三〇代だけれど、日本で若者に、数千万円もの初任給を与えられない。それをする、日本の雇用システムが崩れてしまうから。そうになると、起業するしかないですね。もしくは企業が、別会社を作っていく。

松尾 そうですね。別会社を作る企業も結構あります。しかしできることなら、既存の人事制度を変えたいですね。どこもディープラーニング人材がいらないに困っています。でも給料は変わらず低いま。根本の考え方を変えなければ進みません。適正な給料設定をすれば、人材がなだれ込んでくるはずですよ。

猪瀬 日本人に向いている分野なのに、それを展開するシステムや、企

業の古い統治機構が癌になって、シリコンバレーや中国の深圳から、どんどん引き離されていく。松尾研究室は、一人シリコンバレーをしているようなところがあるけれどもね（笑）。そういう中でG検定は、社会的に認知できる形を作ったことが、大きいですね。

ところで、経済産業省は何かしてくれているの？

松尾 応援はしてくれていますが……本当は資格制度を作るのは、国の仕事ではないかと思えますね。教育を広げる取り組みにも、早く着手して欲しいです。

日本のディープラーニングの研究は、学術的には勝ち目はないですが、そんなことは関係ありません。なぜって、日本の自動車メーカーも電子機器メーカーも、エンジンやトランジスタを発明したわけではない。海外で発明された物を創意工夫して、マーケットを切り開いたわけです。ディープラーニングも同じことをすればいい。

猪瀬 松尾研には、何人ぐらいの研究者がいるんですか。

松尾 正式に所属しているのは三〇名弱です。ただプロジェクトに関わ

っている人は、六〇〜七〇人います。松尾研だけでなく、他大学にも広げる必要がありますし、地方にいる方にも学んでもらえるように、オンライン講座を開設するなど、多角的に進めています。高専の方にも、ディープラーニングに取り組んでもらいたいですね。

猪瀬 技能の習得は、なるべく早い方がいいよね。

松尾 そうなんです。松尾研にも高専出身者がいますが、つべこべいわずに手が動く。しかも、地方に優良なハードウェアメーカーが多くあつて、ディープラーニング人材を欲しがっています。地方それぞれの高専で人材を育てて、地元の企業に貢献するという流れは、ありだと思えます。

そして四〇〜五〇代の方も、時代が大きく変わっているのです。なるべく早く学習し始めた人が得をします。

猪瀬 僕は『昭和16年夏の敗戦』という本を書いているけれど、今の日本はそのときの状況と変わらないような、辛いところがありますね。

松尾 まさに、総力戦研究所のような、今の日本のAIの状況を客観的

に分析し、どこに落ち度やあるいは勝ち目があるのかを考える場所が、必要なのではないかと思えます。

猪瀬 今度は勝つために、総力戦研究所が必要だね。これは国家戦略で

あるべきです。（文中敬称略）
『週間読書人』二〇一八年八月二〇日第三三五号ならびに『WEB読書人』より転載）

講演者紹介

松尾 豊（まつお・ゆたか）

東大大学院工学系研究科総合研究機構 知の構造化センター 技術経営戦略学専攻 准教授。日本の人工知能研究の第一人者。

1975年香川県生まれ。97年東大工学部卒。2002年同大学院博士課程修了。博士（工学）。同年より、産業技術総合研究所研究員。05年より、スタンフォード大客員研究員。07年より現職。02年人工知能学会論文賞、07年情報処理学会 長尾真記念特別賞受賞。著書に『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』『人工知能はなぜ未来を変えるのか』など。

猪瀬直樹（いのせ・なおき）

作家。1946年生まれ。87年『ミカドの肖像』で第18回大宅壮一ノンフィクション賞を受賞。『日本国の研究』で96年度文藝春秋読者賞受賞。2002年小泉首相より道路公団民営化委員に任命される。07年6月東京都副知事に。12年12月に東京都知事に就任。13年12月辞任。15年12月に大阪府市特別顧問に就任。代表作に『昭和16年夏の敗戦』『ペルソナ三島由紀夫伝』『ピカレスク 太宰治伝』。近著に『救出』『戦争・天皇・国家』『民警』『東京の敵』『国民国家のリアリズム』（共著）『明治維新で変わらなかった日本の核心』（共著）。

人工知能に置き換えられない人間の人間としての最後の砦は、五感でしょう。ところが現代社会はその五感を喪失しつつあります。かつてわれわれの暮らしは豊富な香りと音につつまれていました。

室町時代に初代志野宗信が京都で始めた香道は現在まで一度も途切れることなくその伝統を継承しています。今回、志野流香道の蜂谷宗苾・若宗匠より、聞香について実践を交えてご教授いただきます。

香りは聞く、のです。「平家物語」に登場する祇園精舎の鐘の音とは、どのような音色でしょうか。クリスタルボウル奏者・石塚麻実さんによる演奏で、聞香とともに古代の音を味わいます。

特別講演

五感を研ぎ澄ます ——日本古来の香りと音とは？

蜂谷 一枝軒 宗苾氏
志野流香道 21世家元継承者



石塚麻実、クリスタリスト麻実氏
一般社団法人クリスタルボウル・アカデミー・ジャパン代表理事、
クリスタルボウル・アカデミー・ジャパン株式会社代表取締役



司会 猪瀬直樹氏
作家、日本文明研究所所長



日時：2018年8月28日（火）
19時～21時（18時30分開場予定）
会場：日本経済大学東京渋谷キャンパス246ホール
（定員100名）
〒151-0031 東京都渋谷区桜丘町25-17

参加費：2000円（当日、受付にてお支払いください）
参加申込先：下記のサイトよりお申込み下さい。
<http://www.japancivilization.org>
お問い合わせ先 一般財団法人日本文明研究所
TEL 03-5456-8082



伝統工芸品 紹介サイトのご案内

海外サイト：<https://shops.japancivilization.org>
日本国内サイト：<http://japancivil.shop9.makeshop.jp>

所長からのメッセージ

日本文明研究所では、シンポジウムに加え、伝統工芸品の紹介と販売を行っております。日本には現在約1200の伝統工芸品があるといわれますが、それらの生産基盤が衰退しつつあるのが現状です。当研究所は、活動の一環として、全国の貴重な伝統工芸の品々を掘り起こし、永々と受け継がれてきた日本人の匠の技や美を皆様とともに堪能し、育んでいけることを願っております。



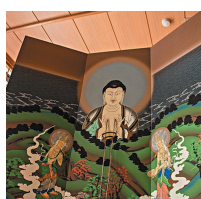
日本竹古風京都籠



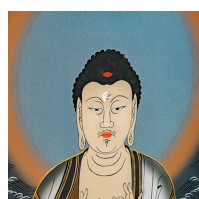
絨（おどし）サムライ鎧



つまみかんざし



山越阿弥陀
（半分タイプ）



山越阿弥陀
（原寸タイプ）



京焼 陶器