

日本の人工知能—将来の展望

Present and Future of AI and Natural Language Processing



国立研究開発法人産業技術総合研究所 フェロー 人工知能研究センター 研究センター長

辻井 潤一

産総研 フェロー、人工知能研究センター 研究センター長、英国マンチェスター大学客員教授、国際計算言語委員会 (ICCL) 委員長、AAMT / Japio 特許翻訳研究会委員長

1 はじめに

我々の研究センター(Artificial Intelligence Research Center - AIRC) は 2015 年 5 月に設立、5 年足らずが経過した。本研究センターは、深層学習の進展がもたらした第 3 期の AI ブームに対応して、日本の AI 研究の中核機関の一つとして機能することを目指して設立された。本稿では、我々の研究センターの現況を報告し、日本の AI 研究の立ち位置とその将来を展望する。

2 3つのリソースの寡占

センターが設立された 5 年前は、AI の研究開発に占める、GAFA と総称される米国の巨大な IT 企業に対する懸念が意識し始められた時期であった。AI 研究に必要なリソースは、データ、計算資源、アルゴリズム (すなわち、それを開発する人材) である (図 1)。この 3 つのリソースが少数の組織に取り込まれるという寡占の状態は、この 5 年間にさらに加速され、日本だけでなく、ヨーロッパ、米国でも大学や公的機関が持つ危惧は、より強くなっている。

サイバー空間にある巨大なデータを収集し、それをビジネスに結び付ける、そして、そのビジネスを通して、データをさらに収集する、特にユーザ行動に関するデータを収集するという彼らのビジネス形態が、大きなデー

タ集合をもとに判断や予測を行う現在の AI 技術を発展させる原動力となってきた。

彼らは、サイバー空間の大きなデータを収集・蓄積・利用するために、巨大なデータセンターを設立・運用していた。この過程で、巨大なデータを管理・運用するために大規模な計算環境を構築し運用する基盤技術の研究開発が、大学や公的な研究機関から GAFA やマイクロソフトなど、ごく少数の企業体へと移動していった。それに伴って、このための計算機技術を研究・開発する人材も、彼らの中に吸収されていった。

この技術の寡占化は、彼らが、ビジネスであげた巨大な財力を使って、データ活用の技術の開発人材、すなわち、AI 人材を高額で雇用し始めたことで、AI 人材へと波及していった。データ活用技術での競争力を高めることが彼らのビジネスの成功につながり、競争力を高める。結果として、AI 技術の人材も財力の豊かな少数の巨大 IT 企業に集積していく。

データ、計算資源、アルゴリズム (すなわち、人材) という、AI の研究開発に必須な 3 つのリソースが、少数の民間企業に集中していく傾向は、将来の AI 技術の研究開発にとって、望ましいものではない。これにどのように対処していくかが、日本の AI 技術を考える際の鍵となる。

- 人工知能の研究、社会実装のためには、単に、アルゴリズム (A) の研究だけではなく、ビッグデータ (B) の収集に係る研究、多量の計算基盤 (C) の構築・活用が不可欠。
- 産総研AIRCでは、2018年8月に運用開始されたABCi、2019年春に本格開始されるAIグローバル拠点と密接に連携しながら、AIの研究開発・社会実装を推進する。

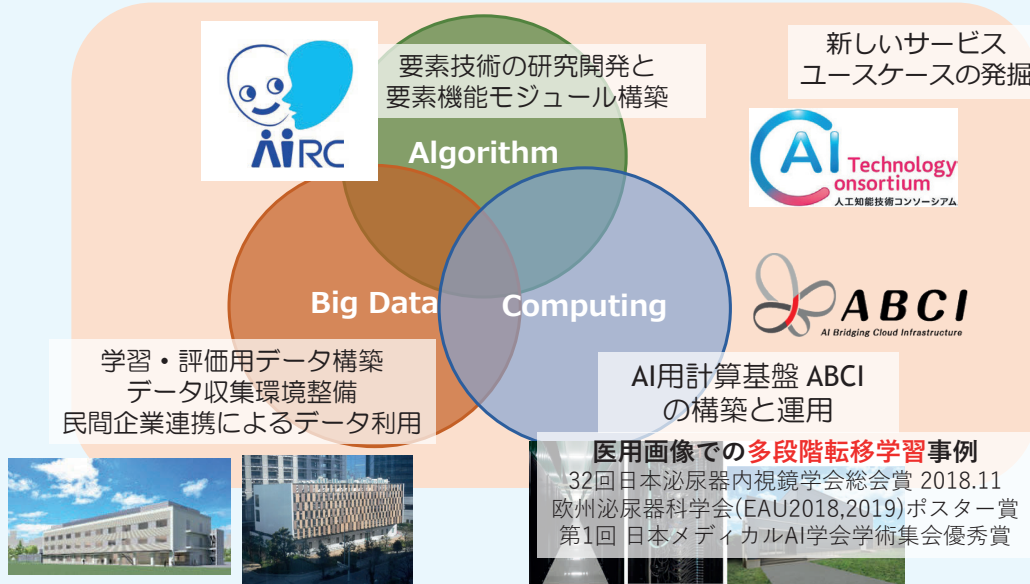


図1 AIRCにおける三位一体の取組み

3 実世界 AI

現在、AI技術の適用範囲は、従来のサイバー空間から大きく拡大し、製造業や創薬などの産業から、人や物の移動、医療・介護まで、人間生活の様々な局面へと広がっている。AI技術は、サイバー空間での技術から、人間生活に直接影響を与える実世界の技術へと拡大し、社会全体に大きな影響を持つ基盤的な技術となりつつある。

AI技術の実世界応用への拡大は、少数の巨大IT企業による技術の寡占状態への危惧をより強めることとなった。医療や介護など、公的な関与が必要な分野、あるいは、人や物の移動のような社会のあり様を決める社会設計の基盤技術が、少数の企業の寡占状態になることは望ましくない。また、AI技術は、製造業など、ほかの産業の競争力を支える技術となりつつある。AI技術の寡占状態は、IT産業以外の既存の産業の在り方にも影響を与える。

ただ、広い分野へのAI技術の展開は、AI技術の寡占化を打ち破る契機ともなる。AI技術は、巨大IT企業のビジネスに組み込まれることで、それらの企業のビジネ

スにあった形で発展した。彼らのビジネスとは別の、それ以外の対象分野にAI技術を適用していく場合には、ビジネスの中でデータを収集し、データを独占するという彼らのモデルは通用しない。たとえば、医療の分野へのAI技術の展開では、病院、医療機関が彼らのビジネスの中で集積してきたデータを活用することが不可欠である。製造業におけるAI適用も同様である。巨大IT企業によるデータの寡占は、AI技術の適用範囲が広がる過程で、自然と解消されていく。

対象分野の広がりや、人材面でも新たな流れを生み出す。医療や製造業へのAI適用では、その分野の専門家、あるいは、その分野が持っている技術とAI技術とを融合するための研究・開発が不可欠である。ここでは、巨大IT企業が抱え込んできた人材だけではない新たなタイプのAI人材が必要となる。

我々のセンターが発足した2015年は、このようなAI技術の多方面への適用が意識し始められた時であった。我々は、この多方面へのAI技術の展開が次のAI技術へと発展していくこと、また、多方面へのAI技術の展開は、必ずしも、巨大なIT企業だけではできず、むしろ、製造業や医療、介護など、多くの分野で成熟した

産業構造を持つ日本が活躍できると考え、この方向に向けての AI 技術の開発を目指し、それを実世界 AI と呼んだ。5 年を経過した現在、この実世界 AI への流れは世界的にも顕著なものになってきている。

4 AI 技術開発の基盤—テキスト処理を例にして

AI 技術を支える基盤として、データと並んで重要な計算資源の問題を考えてみよう。

たとえば、私の専門であるテキスト処理分野では、Google が昨年発表した BERT が評判になっている。BERT は、テキスト処理のための Pre-trained といわれている。

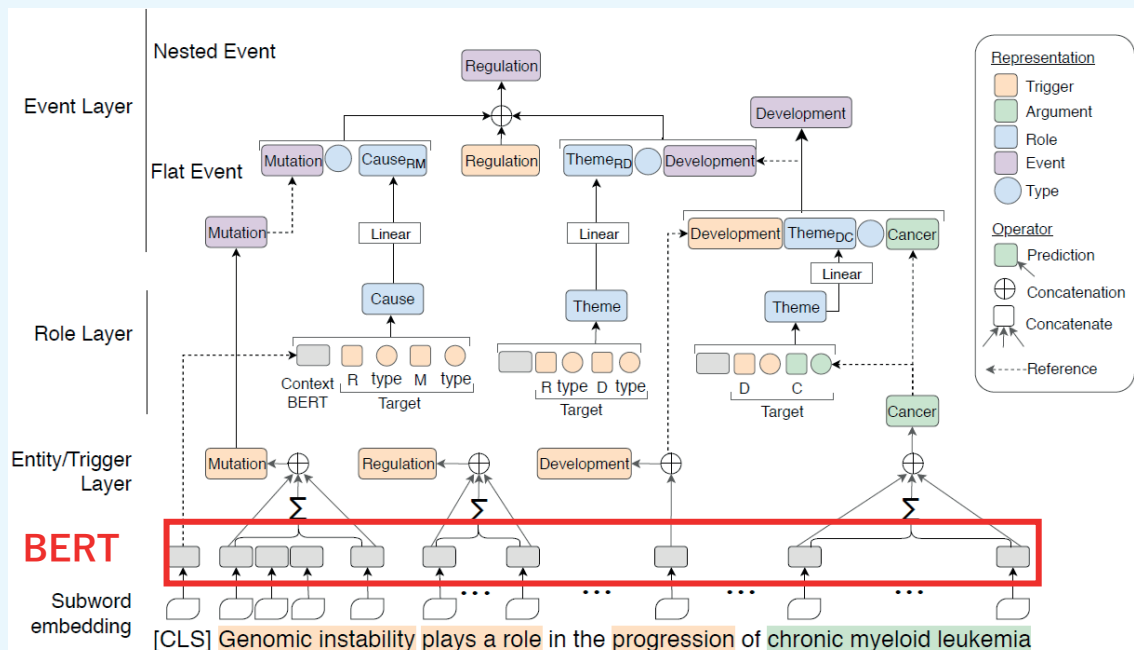
すなわち、この BERT の上に、特定のタスクのためのニューラルネットワークを載せ、BERT を含む全体をその特定タスク用の訓練データを使って再学習する (Fine-Tuning と呼ばれる) ことで、そのタスク用のニューラルネットワークを構築することができる (図 2)。

すなわち、BERT は、いろいろなテキスト処理タスクを実行する個別のシステムをつくるための汎用の基盤モデルで、それを使ういろいろなテキスト処理を行う個

別システムを簡単に作れる、というわけである。実際、BERT を使い、Fine-Tuning されたシステムが、言語処理の分野での標準タスクのすべてで最高水準の性能を示している。現在の言語処理の多くの分野での研究が、(否定するにせよ、肯定するにせよ)、BERT (あるいは、その後継のモデル) を無視しては議論ができなくなっている。

BERT を特定のタスク用に Fine-Tuning するのは、比較的簡単で、多くのグループが BERT を使って研究を行っている。ただ、BERT には Base と Large という 2 つのサイズが異なるモデルがあり、Fine-Tuning をする場合でも、Large を使う場合には、大きな計算資源を必要とし、大学の平均的な研究室では、手に負えない。道具として使うにしても、それなりの計算資源を用意する必要がある。

より深刻な問題は、BERT そのものがブラックボックスになっていることである。Google の作った BERT は、Wikipedia の 33 億語からなる英語テキストを使って作られている。彼らは、Google の 16TPU からなるクラスタを使って数日で学習できたと報告しているが、大学の研究室が持つ GPU を使うだけでは、月のオーダーで時間がかかる。Google の作った BERT は



[1] Makoto Miwa and Sophia Ananiadou. Adaptable, High Recall, Event Extraction System with Minimal Configuration. BMC Bioinformatics, 2015
 [2] Jari Bjorne and Tapio Salakoski. Biomedical event extraction using convolutional neural networks and dependency parsing. BioNLP 2018.

図2 BERT と Fine-Tuning

英語用であるから、日本語や中国語といった他の言語用のBERTは、その言語を処理したいグループが作らなければならない。

現在は、GPUを使った分散学習環境も整備され、TPUを使わなくても、恵まれたGPU環境を整備している大学の研究室（例えば、16基程度のGPUが使える研究室）では、10日程度でモデルを作ることができる。ただ、一つのモデルを作るのに10日程度の時間がかかるとなると、いくつも違ったテキスト集合で少しずつ違ったモデルを作って、BERT的なモデル自体の研究を行うことはできない。Googleは、TPUでBERT自体を訓練するクラウド環境を提供しているが、やはりこの種の研究を行うにはコストがかかりすぎる。実際、現在のところ、BERTやその後継の研究を行うことができる機関は、世界的にもごく少数である。

この種の大規模な研究を推進するためには、我々のような公的な研究機関や大学の研究者が比較的安価に使える大規模な計算機環境を整える必要がある。産総研では、AI研究における計算資源の重要性を認識し、4532基のGPUを持つ計算機環境（ABCI）を昨年度から運用している。

実際、ABCIの4352基のGPUをすべて使うと、Googleが公開時に数日かかると報告したBERTの訓練を20分足らずで実現できる（図3）。

この能力を使って、たとえば、180億語の医療生命分野のテキスト集合でBERTを初めから作り直すと、既存のBERTをFine-Tuningしたときに比べて、医療や生命科学分野のテキスト処理の性能が格段に向上することを確認している（図4）。

ABCIのような大規模な計算環境が比較的安価に使用

BERT Large の実装	環境	学習時間
Tensorflow [1]	16 Cloud TPUs	4日
Tensorflow	64 GPU (ABCI)	4.06日
PyTorch (Apex, FP16)	64 GPU (ABCI)	2.35日

Googleの実験設定と同様の設定でBERTを構築する場合の学習時間
Cloud TPUは4GPU相当
PyTorchがABCI上での計算時間。

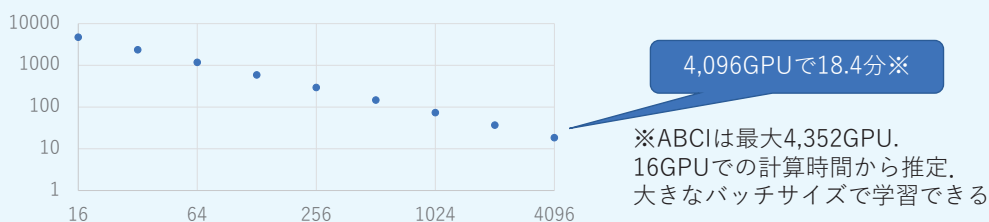


図3 ABCI 上での BERT の推定学習時間

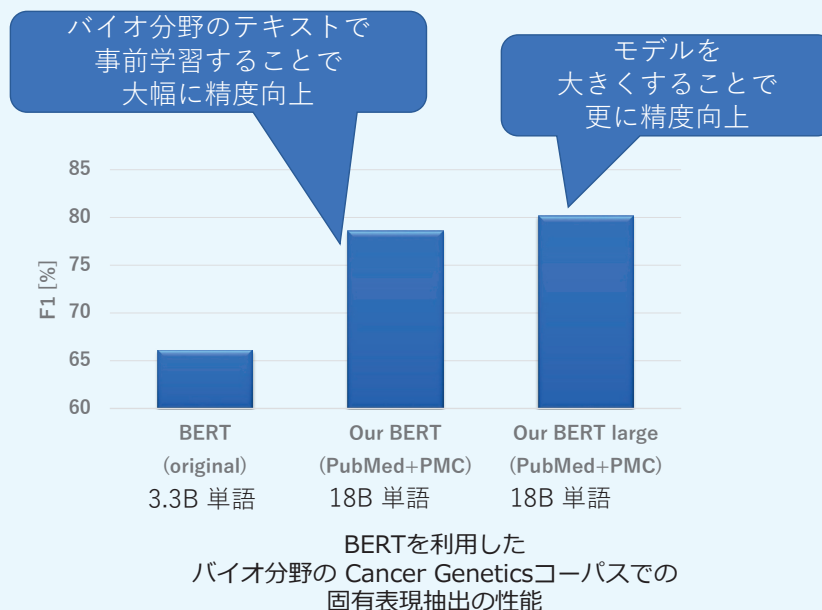


図4 汎用 BERT と医療・生命科学用 BERT

ようになったことで、これまで GAF A のような組織にしかできなかった基盤の技術開発を、日本でも可能にすることができる。米国やヨーロッパ、あるいは、アジア地域でも、同種の計算環境を公的に構築しようとする動きは活発化しており、計算資源の寡占状態も解消していくと考えられる。

5 Pre-Trained モデルと移行学習

前節では、テキスト処理のための Pre-Trained モデル BERT について議論した。このような Pre-Trained モデルを Fine-Tuning することで、個別のタスクを実行するモデルを作るという技術は、他の分野でも成功を収めつつある。

テキスト処理では、論文、特許文書、医療カルテなどから、必要な情報を構造化して取り出す情報抽出 (Information Extraction) という技術がある。この技術は、自由に書かれた、非構造化データのテキストから、その「意味」を取り出し構造化する技術で、意味に基づいた高度な情報検索を実現する基礎技術となっている。

ただ、テキストから取り出したい「意味」は応用ごとに異なる。特定の応用にとっての「意味」がテキスト中にどのように表れるかをシステムに教え込むためには、テキストに意味を付与するアノテーション作業を行い、訓練データを準備する必要がある。

この訓練データをたくさん用意することで、すなわち、模範解答となる見本データを人手でたくさん用意することで、システムの性能を上げる必要があった。この訓練データの作成コストが、情報抽出技術を実際の応用に適用するための最大の障害となっていた。

これに対して、BERT 自体の構築は訓練データを必要としない、Non-Supervised Learning である。例えば、医療・生命系では、オープンアクセスの論文誌が多くあることから、180 億語という大規模なテキスト集合は容易に手に入る。これを使って BERT を作ると、生命医療系のテキストが持つ特徴をとらえた Pre-Trained モデルである BERT ができる。

この医療生命系のテキスト処理用に作られた BERT に、個別の応用のためのニューラルネットワークをのせ、比較的小規模な訓練データで全体を Fine-Tuning することで、個別の応用が必要とする「意味」を取り出すシ

ステムができる。

このように、BERT という大規模なテキストを使って作られた Pre-Trained モデルがあると、それを少数の訓練データで Fine-Tuning するという方法論は、たとえば、特許テキストのような大量テキストが手に入る分野での応用システムの構築を容易にすることが期待できる。

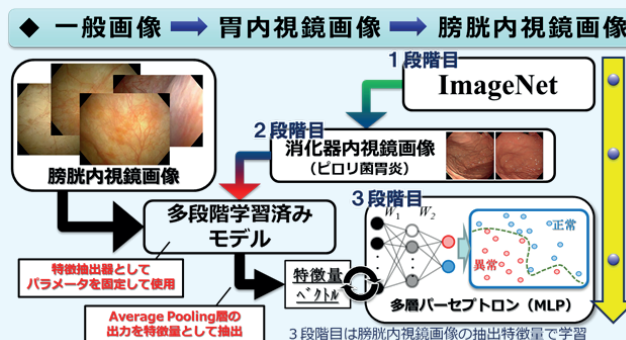
ただ、前節で述べたように、この Pre-Trained モデルは、テキスト処理用に一つのモデル (たとえば、Wikipedia という一般ドメインの BERT) があり、それを Fine-Tune するだけでよいという単純なものではなさそうである。今後、応用ドメインやタスクの種類ごとに有効な Pre-trained モデルの研究が進んでいくと思われる。このような研究を系統的に推進するためには、ABCI のような大規模な計算環境が不可欠となっている。

Pre-Trained モデルの活用は、テキスト処理に限られたことではない。たとえば、イメージ処理では、ImageNet を Pre-Trained モデルとして用いて、医療画像の処理 (例えば、がんの画像検診) を行うことで、良い成果が得られている。ここでも、猫や車などが写っている一般画像を判断する Pre-Trained モデルをまず作り、これに対し医療画像を処理するための Fine-Tuning を行う。

画像を判断するための特徴のようなものが大量の一般的な画像データから学習され、それを Pre-Trained モデルにすると、訓練データの作成にコストがかかる医療画像の判断を行うモデルが、少ない訓練データで可能となる。

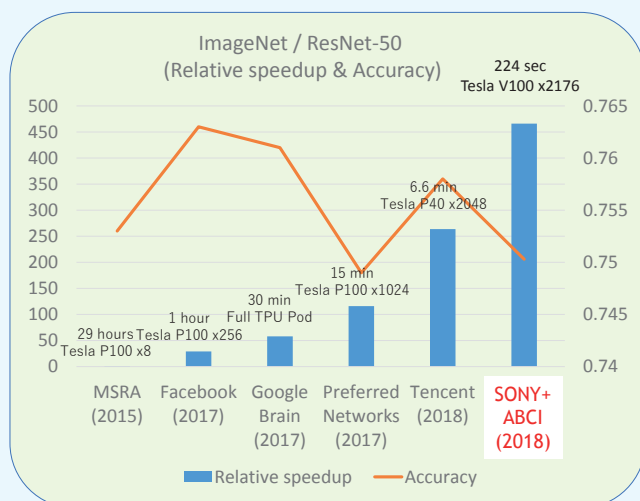
さらに面白いことには、この Pre-Trained モデルを多段階にすることもできる。ImageNet で Pre-Trained モデルをまず構築し、それを比較的大量に作ることができる胃の内部画像からピロリ菌の有無判定を行う訓練データで Fine-Tuning する。その結果を Pre-Trained モデルとして、さらに膀胱内視鏡画像から腫瘍の有無を判定するモデルを作ることができる。この2段階の Pre-Trained モデルが、膀胱内視鏡画像からの腫瘍の有無判定の性能を大きく向上させることができる (図5)。

膀胱内内視鏡での腫瘍有無判定を行うための訓練データは、画像そのものを収集し、それに対する医師に



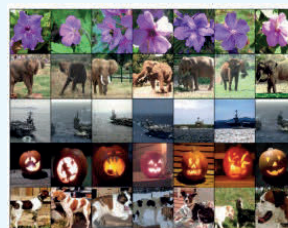
医用画像での多段階転移学習事例
 32回日本泌尿器内視鏡学会総会賞 2018.11
 欧州泌尿器科学会(EAU2018,2019)ポスター賞
 第1回 日本メディカルAI学会学術集会優秀賞

図5 画像処理のための多段階学習



2018年11月
 ABCIの活用により、
**深層学習の学習速度で
 世界最速を達成。**

【ImageNetを使ったタスク】
 120万枚の自然画像を学習して、
 1000クラスの分類を正確に行う



その後 Google によってさらに記録を塗り替えられたものの、
 ABCIにより、Google、Tencentなど先行する技術開発との互角の対抗が可能に

図6 ImageNetの学習時間の推移

よる判断をアノテートする必要があり、少数の訓練データしか用意できない。これに対して、ピロリ菌有無の訓練データは比較的簡単に大量に作れることから、これを第2段階目のPre-Trainedモデル構築に使うことで、内臓内壁に特異的な特徴を捉えたPre-Trainedモデルが構築できる。結果として、膀胱内視鏡の訓練データが少数であっても、高い判別性能を示すというわけである。

今後、画像認識の技術は、医療や製造業のためのAI技術の応用として広く使われていくことになろうが、タ

スクのクラスごとに有効なPre-Trainedモデルが、多数、作られていくことになろう。

ImageNetの判別モデルも、2015年にマイクロソフトが作成した場合には1日以上のかかりかかったが、その後、大規模なGPUクラスタの使用と分散学習のための計算技術が向上し、Sonyと我々産総研とのグループがABCI上で開発したシステムでは数分で学習できるようになっている(図6)。

すでに述べたように、多数のタスクのクラスに依存し

た Pre-Trained モデルが作られていくためには、このモデル構築が短時間で実現できることが不可欠となる。数分でモデルが構築できるといっても、実際には、モデルの最適なアーキテクチャやハイパーパラメータを決定するためには、少しづつ異なるモデルを大量に構築し、最善のパフォーマンスを示すモデルを構築する必要があり、このための計算量は、非常に大きなものとなる。

また、ここでは、2次元画像の認識問題を例に挙げたが、現在、研究は Video の処理へと進んでいる。Video 画像の認識処理では、平面画像の2次元データに時間軸が加わった3次元データとなり、データ量は飛躍的に増大し、構築されるモデルも複雑化する。大規模な計算を行う環境の構築は、今後、ますます重要なものとなっていく。

6 AI 研究の学際性と人材の育成

これまで、AI 研究を支える柱の2つ、データと計算資源を議論してきた。最後の柱である人材について、考えてみよう。

ビッグデータ時代にも、データ・サイエンティストの不足が強く認識され、その育成の重要性が強調された。数理的な理論を理解し、データアナリティクスのソフトウェアを使いこなす能力があること、これはデータ・サイエンティストに必須である。ただ、このような人材の育成だけでは、問題の解決にならないことも強く認識されている。

実際、能力の高いデータ・サイエンティストは、データを生み出す対象分野を理解し、データを対象分野の理解に結び付ける能力が高い人たちである。現在では、医療や経済・経営学など、対象分野の人材にデータ・サイエンティストのための教育を施すことがより有効であるとの認識も広がっている。

もちろん、データを分析する基本的な数理的な道具立てを作り出す人材やそれを計算機ソフトウェアとして実現する人材が必要なことは疑いない。ただ、量的な不足がより深刻なのは、対象分野の知識を同時に兼ね備えた、外向きの人材である。

人工知能においても、この傾向は同様である。いや、データ・サイエンティストが目指した対象分野の分析や理解から一歩進んで、対象分野での課題解決を目指す人

工知能の研究開発では、この傾向はさらに強い。

たとえば、がんの治療薬を作る AI 技術の研究が、AI 研究者・技術者だけでできないことは明らかであろう。また、製造業で使われる知的なロボットの研究でも、製造業が持つ既存の技術と AI 技術の融合が不可欠であり、AI 研究者だけで新技術ができるわけではない。実世界 AI では、実世界を対象に問題解決を行ってきた、その分野の研究者や技術者との協働が不可欠であり、これらの人材は必ずしも巨大 IT 企業が寡占的に集まっているわけではない。

これら他業種で成熟した組織体とそこでの人材を豊富に持つ日本は、AI 研究者が彼らとの協働関係を構築していく上でも高い可能性を持っている。

7 人間との協働

AI と人間は、いずれも、対象分野についてのモデルを内部にもち、そのモデルをもとに判断し、行動する。内部のモデルに基づき、自律的に知的な判断や行動をとることが、知的とされる所以である。

ただ、この2つの知性体が内部に作り出す対象のモデルは、かなり違ったものである。データに基づく現在の AI は、対象に関する大量のデータを収集し、そこから帰納することでモデルを構築する。大量データから、そこに潜在する規則性を帰納する能力においては、AI は人間の能力をはるかに凌駕している。

これに対して、人間は、限定された、少ないデータからモデルを構築できる。この過程には、幼少期からの教育によって習得された、対象の背後にあるメカニズムについての理解がある。たとえば、医療従事者は、幼少期からの教育の基盤の上に医学に関する教育をうけ、それをもとに対象を理解し内部的なモデルを構成している。データ化されていない、稀な出来事に対しても、人間が対応できたり、少ないデータからも適切なモデルが構築できたりするのは、演繹的な知識に助けられているからである(図7)。

このような異なった特質をもった2つの知性体、人間と AI とが協働して課題を解決していく枠組みやそのための技術を開発して行くことが、今後の AI 技術の研究方向である。現在喧伝されている AI をホワイトボックス化する技術、いわゆる、説明できる AI の技術開発は、

- 人間は言語空間、ルールに基づく知識処理が得意。
- AI（サイバーの情報処理）はRaw Dataのビッグデータ処理が得意。
- 人間の思考もAIも具象、抽象、学習、シミュレーションにより価値のある情報を生成する。
- 具象、抽象の間をつなぐStructureは人間の思考もAI（機械学習）も不明点が多い。
- XAIはAIにおける、この構造を明らかにする。人間、AIの構造化をつないで、片方の進化を他方の進化につなげるのが共進化AI。

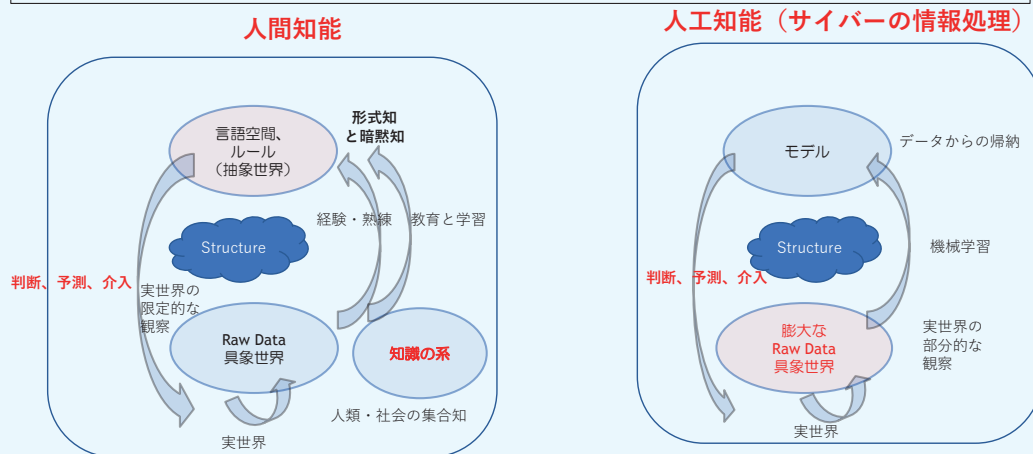


図7 人間知能とAI

このための必須な技術である。

また、帰納を中心とするAIとは異なった技術の流れもある。例えば、製造業でのデジタル・ツインの技術は、演繹的なモデル化手法であるシミュレーション技術を使うことで、サイバー空間に対象のモデルを作り、これを操作することで設計、開発を加速する技術であった。この演繹的な手法であるシミュレーション技術とデータからの帰納的なAI手法とを結びつけようという研究も、今後、さらに活発化していくものと期待している。

デジタル・ツインの技術は、巨大IT企業が開発してきたというよりも、製造業のエンジニアが中心となって作り上げてきた技術であり、個々の業種ごとに多様な技術の展開を遂げている。これらの技術とAI技術の統合も、日本のAI技術の進展を考えていく上で重要な課題で、日本がもつ豊かな（AI人材以外の）人材を活用していく道であろう。

8 データの積極的な獲得

AI技術が開花したのは、それに先行するビッグデータのブームがあり、大量データを取り扱うための計算機技術や機械学習の研究が進んだことが、その素地となった。

データアナリティクスというデータ分析ツールの開発に代表されるように、ビッグデータの時代はデータを分析することで、人間が対象を理解することを助けるという傾向が強かった。

これに対し、AIは、対象の内部モデルを構築することで、判断や行動を行うことができる自律性の高いシステムへと向かう。この枠組みでは、対象の人間による分析を助けるというよりも、ある特定のタスクを自律的に行うシステムを作るためにデータをどのように使うか、あるいは、より積極的に、あるタスクを行う自律システムを作り出すためには、どのようなデータを取得するのが良いか、という問題設定を行う。

大きなデータがあるからそれを分析するという受動的な立場から、あるタスクを行うAIシステムを構築するためにはどのようなデータを取るべきか、という積極的な立場への変化がある。

AIの研究開発では、あるタスクを行うAIシステムを開発したいが、データが十分にそろっていない、そのためデータをとることから始めなければならない、という嘆きがしばしば聞かれる。

このことは、決して悲観的な嘆きではなくて、次のステップのAI技術の研究開発には、IoTやスマートセンサーの活用を問題解決に結び付けていくこと、データ取

得により積極的に関与することで、トップダウン的に既存の産業形態を変革していく態度が必要である、ということであろう。

このような背景から、産総研では、AIによる問題解決とデータ取得を並行的に行うために、ロボットのいる生産環境、ロボットによる生命科学実験環境、コンビニでのマテリアル・ハンドリング環境を備えた実験棟を構築し、その運用を開始している（図8）。

9 おわりに

本稿では、設立5年になる産業技術総合研究所・人工知能研究センターの現在の研究開発方向を中心に、日本が目指すべきと考えるAIについて議論した。

米国の巨大IT企業、さらには、中国の国家をあげた

AIの育成など、AIの研究開発競争はさらに激化している。AI技術は、サイバー空間から、実世界の幅広い活動に影響をもつ、社会の競争力を左右する基盤的な技術としての役割を強めている。

日本だけでなく、ドイツ・フランス・英国、あるいは、台湾・シンガポール・インドなどがこの競争が激化する技術開発にどのような戦略で臨むか、模索を重ねている。

個人的には、AI技術が社会の幅広い分野に影響を与える技術へと進化していくことは、社会の多くの局面で成熟したシステムを持つ日本にとっては、貢献していく好機が訪れていると考えている。そのためには、新たなAI技術を開発していくための資源（人材、計算機環境、データ）を日本としていかに確保していくかの戦略を持つ必要がある。

- 労働生産人口低下で影響を大きく受ける**労働集約型の産業**として「工場」「物流」「創業」の3つの環境を模擬環境として構築
- 模擬環境からAIの学習データ取得するため、各模擬環境のサイバーフィジカルシステムを構築
- 模擬環境をテストベッドとして活用し、ここで開発された仕組みを実際の現場へ展開

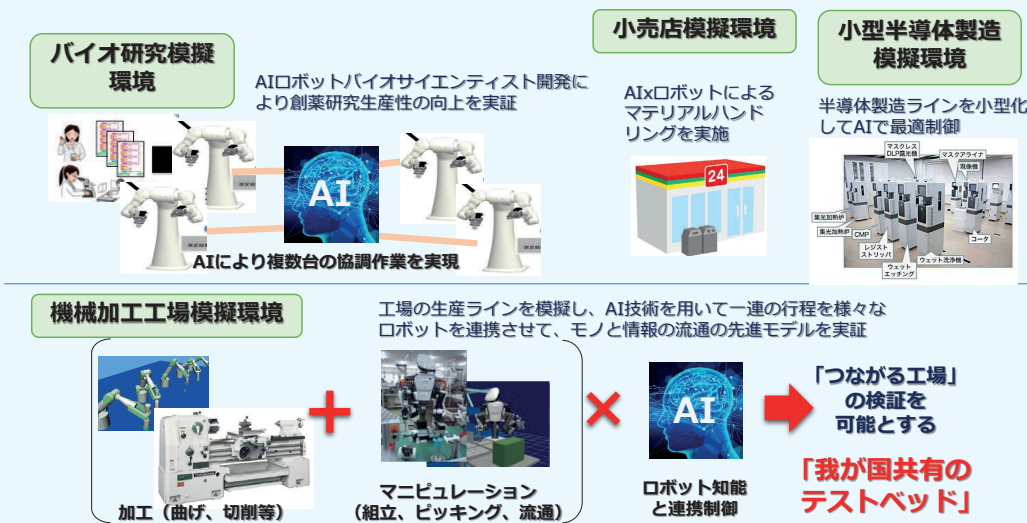


図8 データを積極的に取得する実験棟とその模擬環境



特集

特許情報分野におけるAI活用のおすすめ