

[第 12 回シンポジウム 基調講演]

AI 社会実装における要件と AI 技術の擦り合わせ

慶応義塾大学 理工学部管理工学科 教授,
人工知能ビッグデータ研究開発センター長

山口 高平 氏

この記事は、情報システム学会 第 12 回シンポジウム (2019 年 5 月 25 日) における基調講演の口述内容をまとめたものです。

■はじめに

みなさん、こんにちは。明日から会長を務める、慶應義塾大学理工学部の山口です。

AI が今、ディープラーニングを中心に第 3 次 AI ブームを迎え、多くの会社からよく相談を受けますが、AI 利用に関する要件が曖昧なまま話が進んでいるなあと感じています。すなわち、AI 技術は細分化されており、それぞれの AI 技術はかなり異なるにも関わらず、AI という言葉で一括りにしているために、話しが曖昧になっているということです。例えば、機械学習とディープラーニングを比べてみても、うまく動かすための条件は異なるのですが、そういう差異は認識されないで、話しが進んでいます。また、データをたくさん持っていれば、AI を実行できるという勘違いがあります。日本はこのような状況ですが、米国では、そのような話がだいぶ少なくなってきたと米国の知人から聞いています。我が国のこういう状況において、情報システム学会では、AI システムの要求仕様、要件、開発技術を擦り合わせていくことが、大きなポイントになっていこうかと思えます。ということで、今日は、そのような視点から、AI 技術の日進月歩も踏まえながら、AI 技術の進展

とユーザの要求を擦り合わせていく重要性について、私の経験を交えてお話しできればと思います。その後、皆さんからいろいろなお意見、質問を承れればと思います。

■中国の進出

まず、当学会のシンポジウム講演は、私は 3 回目になります。2010 年には、サービスサイエンスの話をしました。3 年前の 2016 年は「AI がもたらす新しい社会」について講演し、学会誌の 12 巻 2 号で読むことができます。AI 技術の進展につきましては、この時かなりお話ししましたが、この 3 年間でまた状況が大きく変わりました。それは、中国の AI への進出です。3 年前の中国の AI は注目すべきことはあまりなかったのですが、この 3 年間で AI の社会実装が急速に進みました。例えば自動運転のスタートアップが 10 社近くできて、自動運転技術について、日本の車メーカーが追い越されてしまったという厳しい現実があります。国が投資する予算も、日本は去年が 700 億円で今年が 1200 億円となってすごく増えたのですが、中国の AI 研究開発投資額は 1 兆円規模になっています。桁が違うわけですね。中国の AI 社会実装は米国でさえも一

目おいています。後で、このあたりにも触れたいと思います。

■エンドユーザーの要求は変わる

要件と技術の擦り合わせにつきましては、私より聴衆の皆さんのほうがお詳しいと思うのですが、情報システムの典型的な開発プロセスとしてウォーターフォール、あるいはV字型モデルがあります。要求分析から要件定義、基本設計から詳細設計、そして実装し、レビューしながら単体、結合、総合テストをしていくという、V字型モデルが基幹系システムならまだ利用されているかと思います。ただ2000年になってWebアプリケーションの開発になると、動かしてみないとわからない、特にソフトウェアあるいはロボットなどは動かしてみると、「いや、これは違うなあ」とエンドユーザーは要求を目まぐるしく変えてくるわけです。要求がどんどん変わるので、それに対応する開発方法論としてアジャイル開発方法論が出てきました。基本的機能を実装し、まずは動かしてエンドユーザーに見せる。そうすると、「あ、ちょっと違うな、ここはこうした方がいいなあ」というように新たな要求が出てきますので、それを反映してまたすぐに動かす、こういう一連の開発プロセスを小さな単位で繰り返していきます。そういうプロセスがアジャイル開発で、Web系の人たちはこれを中心にして開発をしている状況かと思います。

そうは言っても、エンドユーザーの要求がとにかく変わるということに対応するのはコストがかかりますので、何とかしたいという思いはあるかと思います。ソフトウ

ェアの開発指標としてのメトリクスとしまして、機能性、信頼性、ユーザビリティ、効率性、保守性、ポータビリティのようなメトリクスが、ソフトウェア工学の分野において既に確立されています。これらのメトリクスからソフトウェアを評価する、ただ、これらで評価してこの時点で良かったと思っても、要求が変わればまた一からやり直しになります。

皆さん、この漫画を見たことがありますか？これは20年ぐらい前の風刺漫画です。エンドユーザーがこのようなブランコを作りたいと思っていて、これが要求仕様になります。実際に作るとこうなり、プロジェクトリーダーが理解した範囲内のブランコがこれで、でもこれで揺らすと木にぶつかってしまいますね。そのためアナリストが「これではぶつかるから切る」と考えます。でも切ってしまうと、支えている部分が非常に細いので壊れてしまう、とプログラムのコーダーは考え「ヒモに板を縛り付ける」と考え、営業の人は「板ではなく、豪華な椅子にしましょう」というような感じで、開発途中で要件がどんどん変わってしまうわけです。こういう風刺漫画が20年ぐらい前にありました。結局、いろいろなプロジェクトのドキュメントは影ぐらいしか残っていないとか、顧客の要求を聞いていたら、ジェットコースターみたいにドーンと上がったたりドーンと下がったりするシステムになってしまった、というようなことが起こります。顧客は最初こんなことを言っていたけれど、実はインタビューをしていくと、まあ、実際のコストなどを考えたらこれでいいか、ということもあるのですが、コミ

コミュニケーションがうまくいかなくて、当初の要件と全く異なったシステムが開発されてしまうことが起こります。要求仕様を最初に定めるというのはソフトウェアに限らず難しいわけですが、その維持方法が、大きな課題になっているわけです。

■SERVQUAL

一方、サービスと言う観点からソフトウェアを捉えるという研究があり、これはソフトウェア工学というよりマーケティングに関連しており、SERVQUAL というこういうメトリクスが 30 年ほど前に提唱されました。プラスマン (Prasuraman) という人達が提唱して、提供されるサービスで先程のようなエンジニアリング的なメトリクスではなくて、信頼性、反応性、確実性、共感性、有形性のような、人中心のメトリクスが整理されました。例えばホテルのサービスだったら、これらのメトリクスをどのように具体化すればいいのか、マーケティング分野のサービス工学とかサービスマネジメントという分野で研究が続けられています。例えば、有形性というメトリクスは最新の設備、外的な見かけを意味します。見かけがいい、あるいは、従業員の身なりや態度が洗練されてと良い感じを受けるとか、オーダーしたら確実に間違えずに持ってくる信頼性とか、レスポンスが非常に丁寧である反応性とか、確実性とか、共感できるとか、先ほどのソフトウェア工学のメトリクスからするとすごく主観的で感性的なのですが、サービスレベルの評価というのはそういうものだ、ということで、マーケティングでは研究されています。

■大括りの AI

これから AI の話をしていきますけれど、AI には色々な AI があります。技術論でいうと 200 程度の技術があります。私は AI 研究に 40 年程度携わり、一つ一つの AI 要素技術を研究してきました。しかし、メディアで言及される AI は、これらの AI 技術を一括りにして、中身の AI 技術についてはほとんど言及されません。メディアに出演し、詳細な AI 技術に言及しようとしたら、「いや、そんなことは細かくて視聴者はわからないから言う必要はありません」とか言われます。大括りの AI です。大括りで AI の話を進めようとするので、それが誤解を生んでいると言えます。

■知識系 AI

AI を大きく分けると知識系 AI とデータに対して機械学習するデータ系 AI に分かれます。このスライドの左側が知識系 AI です。1980 年代の AI は、シニアの方が多いのでエキスパートシステムという言葉が記憶されている方もおられるかもしれませんが、ルールベースで専門家の知識を作って動かすという AI でした。今のディープラーニングが人を超えるという論調と同じで、if~then ルールで専門家の知識をすべて表現すれば、専門家を超えるシステムも開発可能である、というようなことが言われました。でも、実際にはそのようなことは無理でした。それは、人の知識は本当に膨大でかつ深いからです。知識の一部分を if~then ルールで記述したところで、少しは賢く動くけれども、状況が変わると、途端に処理できません。そういう現実が分かっ

てくると、エキスパートシステムは、だんだん失望され、言葉の意味理解、常識推論をやらないと駄目だね、ということになり、概念の意味をコンピュータに理解させるオントロジーの研究が始まり、今なお、研究は続いています。

対話システム、AIスピーカーが出てきて、一見対話できているように思えますけれど、それは過去の対話データから、こういう事を言ったらこういう事を言い返したらよい、という表層的なパターンで対応しているだけで、言葉の意味を理解した反応ではありません。それを実現するには、オントロジーと呼ばれる、言葉と言葉の意味ネットワークみたいなものが、米国では1980年代半ばから Cyc (サイク) と呼ばれる研究が始まって、今は100万以上の概念群が定義されています。今なお国防省から研究資金が出資され開発が続いており、非常に息の長いAIプロジェクトになっています。さらに、このオントロジーとディープラーニングを連携するAI研究も始まりました。ディープラーニングは、精度は高いですが、なぜそうなったのかという質問に答えられないため、説明できるAIがキーワードになってきました。AIは説明機能がないので、AIは人とインタラクションが取れない。「なぜこういう判断をするんだ？」と尋ねたら、「これこれこうだから、こうなんだ」というふうにAIシステムが人に答えてほしいわけです。そういうAIをXAI(eXplainable AI)といい、長年開発されてきたオントロジーの利用が検討され始めています。

IBMが知識系AIシステムでは大きな研究成果をあげてきていて、後でご紹介する

ワトソンが2011年にクイズ番組のグランドチャンピオンに勝って、2014年から事業化しています。ワトソンの後継として、プロジェクトディベーターという研究も始まり、人間と討論をして、全米地区チャンピオンだったら、その討論に勝てるようになりました。AIが人を論破するわけです。人を論破するAIには不安感を抱く人もいるかもしれませんが、IBM研究所ではすでに開発されている、これが知識系AIの最前線です。

■データ系AI

一方データ系、機械学習のAIというのは、このエキスパートシステムのブームの時に、バックプロパゲーションに基づくニューラルネットワークが登場し、これは、入力層があつて、中間層が一層または二層あつて、出力層がある、3~4層のニューラルネットワークで、従来のニューラルネットワーク(パーセプトロン)より性能が高く、ブームになりました。ただ5層位になると、訓練データに過度に適応して、新しいデータを認識処理できないというオーバーフィッティング問題、また、30年前はまだコンピュータの速度が遅かったので、学習時間に1週間かかるとか、計算時間問題が出てきました。1990年代に入り、数理統計理論に基づく機械学習として、ベイジアンネットワーク、サポートベクターマシン、ランダムフォレストなどが現れ、これがこの3~4層のニューラルネットワークよりも精度が高く、かつ計算時間も短くて済んだので、1990年代~2000年代は、この3~4層のニューラルネットワークは衰退しました。この時期は、高い精度が出たSVM(サポート

ベクターマシン)がよく使われ、通常、こういう曲線で分離することが、次元（ディメンジョン）を変えると直線分離になるとかいう特徴があって、計算も速いという特徴がありました。

その後に出てきたのが、今のディープラーニングです。ディープラーニングの原型は、1989年にルカンという人が、今フェイスブックのAI研究所の所長ですけれども、その人が提案しました。けれど、コンピュータが遅いために精度が悪かったのですが2010年頃のコンピュータですと性能が高くなって精度も良くなってきました。2010年頃、10層程度のニューラルネットワークでも動くようになり、現在は、後で説明しますが、100層以上でも動くようになっています。ディープラーニングもいくつかに分かれ、画像認識系、言語処理系、生成系（似顔絵を簡単に作るとかは生成系と言います）という3タイプに分かれ、多くの応用が成されています。そういう時代が第3次AIブームと言えます。

一方、知識系AIは、知識ベース構築コストの問題があり、ブームのレベルには至っていません。しかし、先ほど言いましたように、どうしてそうなったの？と聞いてもディープラーニングは一切答えません。細かい計算過程を見せられても人間には全然わからないため、この2種類のAIを連携したXAIが次世代AIの一つである考えられ、米国では11大学がこのテーマに取り組んでいます。

■知識系 AI「ワトソン」

IBM 研究所により開発された知識系 AI,

ワトソンが、ジェパティというアメリカのクイズ番組に出演しました。両端のお二方は人間のグランドチャンピオンです。真ん中がワトソンで AI コンピュータに繋がっていて、2 の二人に挑戦し、ワトソンが見事に勝利しました。IBM では、社内ベンチャーがその年から始まり、2014 年からワトソンは事業化され、実際にビジネスで使われるようになりました。これは日経の記事です。2017 年には売上高 1 兆円程度を達成しました。IBM の全世界での売上高が 10 兆円程度ですから、10 分の 1 程度をワトソンが稼いでいるわけです。世界で一番稼いでいる AI は何ですか、と聞かれば、このワトソンということになります。

■知識系 AI「プロジェクトディベーター」

IBM 研究所から、プロジェクトディベーターという、討論する AI が昨年登場しました。例えば「宇宙資源の探索研究に国は助成すべきか？」という議論テーマを与え、プロジェクトディベーターは肯定側、米国地区チャンピオンが反対側に立ち、議論で対決するわけです。4 分間双方が自分の意見の正しさが主張し、その後「君のさっきの主張のこの部分は、～という理由で、おかしいだろう」などと反論を交え、最後の 2 分間で自分の意見を取りまとめます。人は反対側なので、「宇宙探査にそんな国の税金を使うべきではない」という立場で挑みます。ディベーターは賛成側なので「宇宙探査はこのようなことに貢献して、人類に大きな恩恵をもたらすから、国がもっと助成すべきだ」と、ネット検索結果などを利

用して主張し、その後、意見の弱点を突きながら、反論するわけです。その討論の様子を数百人の聴衆が聞き、どちらの弁が良かったかと尋ねた結果、人ではなく AI (ディベーター) が勝利してしまったわけです。宇宙探査以外にも「遠隔医療をもっと推進すべきである」というテーマで討論したところ、やはり、ディベーターが勝ってしまいました。ただ全米チャンピオンが登場した時は、ディベーターは勝つことができませんでした。さすが、全米チャンピオンという感じですが、地区チャンピオンと全米チャンピオンの中間的な能力がディベーターにはあるよね、というのが現状の評価です。AI 技術としては、ネットから自分の主張を裏付けるような情報検索として意味検索。これは、通常の検索エンジンよりも高度な検索エンジンです。ときどきジョークも言うようです。また、相手方が反論してきた時、その背景にあるその反論の意図を推定することもできます。知識基盤としては、オントロジーのデータ部分であるナレッジグラフを基盤にしています。現在、この程度までしか技術的には公開されていませんが、最高レベルの知識系 AI といえます。

■データ系 AI (画像認識)

このスライドはディープラーニングの説明で NVIDIA の研究者がよく使うもので、入力層の各ノードが各画素に対応するデータになります。オーディとかスカイラインのような車種を認識するディープラーニングを考えた時、一番右側の出力層の各ノードが車種になります。画像が 32×32 だった

ら 1024 の入力ノードにそれぞれの画素の色のデータが与えられ、最初の認識では、間違った認識結果になります。オーディなのにスカイラインなどと誤認識します。その後、バックプロパゲーションにより、ノード間のリンクの重み、0~1 の値が割り当てられていますが、認識結果が正解になるまでに少しずつ重みを変えていきます。そうするとこの辺りの中間層に、エッジを認識するノード、別のこの辺りの中間層だったら、車の部位、例えばタイヤとかドアの部分とかを認識するノードが出現し、最後の出力層で、車種が正しく判別されていくように学習が進んでいきます。すなわち、画像系のディープラーニングでは、低次元の特徴量から高次元の特徴量までの認識がこの中間層に段階的に出現するということが分かってきました。

イメージネットという 1000 種類ぐらいのカテゴリーを認識するコンテスト、競技会があります。この棒グラフの縦軸はエラー率です。2011 年までは、サポートベクターマシンのエラー率が一番低く、2011 年には 25.8% でしたから、4 枚の画像を見せられたら 1 枚は間違えるという、精度でした。人間の精度とは言いますと、だまし絵みたいな写真もあるので、エラー率は 0% にはなりません、5% 程度です。100 枚見せられたら 5 枚程度間違えるという認識率です。しかし、2012 年、ディープラーニングのチームが参戦した結果、25.8% のエラー率が 10% 程度下がりました。このディープラーニングの構造は 8 層程度でした画、2014 年 Google ネットというディープラーニングは 20 層程度になり、エラー率が 6% 程度にな

り、人のエラー率 5%にかなり近づきました。さらに ResNet という、バックプロパゲーションにショートカットを加えた、新しい学習法が登場し、150 層程度のディープラーニングが出てきました。このあと、ショートカットをコンパクトにして 50~100 層程度の SeNet が現在普及している状況です。利用されている方もいるかもしれませんがね。この SeNet のエラー率は 2.3%ですから、人間の 5%より精度が高く、画像物体認識について、SeNet が人間の眼を超えたと言えるわけです。シンギュラリティという言葉がよく出てきますが、ある画像物体認識のような特定タスクに限定すれば、シンギュラリティは既に起こっているとも言えます。このため、画像検査の分野では、人よりディープラーニングを使う状況が増えてきました。その一例が、この NHK の E テレの番組です。ちょっとご覧ください。

<ビデオ再生 以下の一部>

医療画像診断（大腸がん発見）

NHK E テレ AI 超入門 2017.11

最後のナレーションは少し煽っていましたが、お医者さんがいくら頑張っても初期のがん画像診断では 80%の精度ですが、AI は 98%です。これは NEC と国立がん研究センターの共同研究で、大腸がん検査から始まって胃がん検査なども実験し、すべて 98%程度の精度です。もう患者は、人のお医者さんではなく、ディープラーニングに検査してもらうべき時代になったと言えます。ただ、検査後の治療法の決定については、まだまだ、お医者さんの方が AI より優れています。でも画像検査については AI

に診てもらった方がいいわけです。結局画素単位で、ここが少し赤くなってこっちが暗くなっている、その関係を見出すので、人はそのような細かい関係性を認識できないので、ディープラーニングの精度が高くなるわけです。

■自動運転

自動運転、これがディープラーニングの最大のアプリケーションと言われていました。Google からスピナウトしたウェイモという会社は、昨年 12 月に有料自動運転タクシーのサービスを始めました。8km 走って 770 円、まだ事故は起こっていません。このサービスを始めるまでに 1,600 万 km の走行実験をしています。ディープラーニングがかなり使われているのですが、大規模な走行実験データが必要です。走行実験をして、屋根に乗ったライダーが、150m 先の 360°全方位、何が居るか・動いているかを認識するわけです。人間にそのようなことはできるはずがありません。150m 先の信号が今、青から黄色になったとか、そういうレベルですべてを認識するわけです。その認識結果のもとで、今スピードを少し上げようとか、車線変更をした方がいいとか、判断をするのです。同じ制限速度の範囲内で車線変更したら抜けないので、ちょっとスピードを上げて、60km だったら 70km, 80km にぐっと上げて、追い越すときだけスピードアップして、また自分の走行車線にきたら普通の 60km に戻る、そのようなレベルまで、ソフトウェアとして実現されています。

ただ問題もあります。大雨が降ってきた

ら雨を小石と間違え、小石が振ってきたからということで急に止まったりします。日差しにも弱いです。人もそうですけれども、AIシステムは日差しにすごく弱いです。それから、透明な物体に弱いです。あとでお見せしますが、私の研究室でロボット喫茶店という実験をやっていますが、ある時、透明のボードを置いていたら、ペッパーがぶつかっていくのです。いや、そこはパネルがあるから駄目よ、と言っているのに、何もないと思って透明のパーティションにぶつかっていったりします。透明、日差し、大雨とか大雪、それらは、ディープラーニングでも、今のところ正しく認識できません。人間だったらそういう状況になったら、「あ、危ないな」と思ってスピードを落とすとかします。その臨機応変さがまだなくて、Google はそれが今研究課題だと、すでにサービスを開始しているのに研究課題は山積していると、明言しています。

自動運転の宅配、これはクローガーというアメリカのスーパーが、ニューロというスタートアップ企業、このニューロというのはアメリカのスタートアップで自動運転を専門にやっている会社ですが、そこと連携してピザとか色々な食品の配達を自動運転により有料で実施しています。

■中国の躍進

最初に説明しましたように、中国のAI社会実装は3年程前にはほとんど進んでいませんでしたが、習近平国家主席が2030年に中国がAI世界一になるという目標を掲げて以来、この3年間、中国政府が多くの

AI研究開発資金を投入して、スタートアップ(ベンチャー)を育成し、AI社会実装を進めています。これから、2019年1月に放送されたNHKスペシャルをご覧ください。

<ビデオ再生 以下の一部>

NHKスペシャル:アメリカ VS. 中国
“未来の覇権”争いが始まった、
2019.01.19

このスタートアップは1年半前に起業した会社です。たった1年半前です。彼は元々アップルで自動運転のエンジニアでしたので、米国は「中国は米国から技術を盗んでいる」というような批判をしています。GAFAで勤めていた中国人が、中国へ帰ってスタートアップを立ちあげて、米国のライバル社になろうとしているのです。自動運転ソフトウェアを開発する場合、走行実験データが非常に重要です。人の流れを表す人流データが重要です。例えば、交差点で大人と子供が一緒にいたら子供はそれほど急に動かないけれど、子供同士だったらパッと飛び出してくるとか、そういう人流データがないと、自動運転ソフトウェアが賢くなりません。その人流データを中国政府が取得し、中国のスタートアップに提供しているのです。7億人の人流データをいたるところ設置されたテレビカメラから取得し、スタートアップに提供するわけです。これが、国家資本主義といわれる実態です。また、NHKの放送ではGoogleとテスラを同列に並べていましたけれど、これは正しい報道ではなく、圧倒的にGoogleの方がテスラより優れています。テスラは今まで自

動運転により 3 人の死者（すべて運転手）が出ています。現在米国では、自動運転ソフトウェアは、安全性により格付けされる時代になりつつあり、カリフォルニア州道路局からそのレポートが出ています。

■カリフォルニア州のレポート

「自動運転を開始して、危なくなれば自動運転から人に交代する」までの平均距離を解除(disengagement)距離と呼び、カリフォルニア州道路局が毎年報告しています。Google が一番安全で、解除距離は約 18,000km、テスラはトップ 30 位までに入っておらず、ウーバーは 3km ぐらいです。2 位は GM ですが、GM が買収したオートクルージングという米国のスタートアップが自動運転ソフトウェアを開発し、解除距離は約 8,000km です。3 位、4 位までが米国の IT 企業で、NHK で放送された中国のスタートアップ Roadstar.ai というのは 10 位で、解除距離は 300km ぐらいです。日本企業では、日産が 6 位で解除距離はやはり 300km 程度です。6 位から 10 位までは、解除距離は 300km 程度で、ほとんど差がありません。でも 1 位は 18,000km なので、1 位と 6 位は、安全性に 60 倍程度の差があるとも読めるわけです。トップ 10 のうち 1 位～4 位と 9 位が米国 IT 企業、5 位、7 位、10 位に Roadstar.ai、解除距離 2,000km の Pony.AI、バイドゥ、この 3 社が中国です。ただし、人流データを提供する中国政府が後ろ盾になって支援しています。6 位に日本の日産が入っていますが解除距離は 300km で、トップ 5 と比べると大きな差がつけられており、熾烈な競争が自動運転の世界では展開

されています。

しかし残念ながら、自動運転による死者も出てきました。テスラの自動運転では、すでに 3 人の死者（運転手）が出ており、ウーバーも 1 人の死者が出ています。特に、ウーバーの場合は、死者が運転手ではなく歩行者で、米国で昨年 3 月、ウーバーの自動運転車が歩行者を跳ねてしまいました。これは大問題だということになり、米国では、従来、自動運転実証実験に寛大だったので、ウーバーによる死亡事故以来、多くの法律家が関与し始め、法律により自動運転規制の検討が進んでいます。

■AI を大括りする課題

アジャイル開発方法論ということで、ユーザが AI に過度の期待を抱き、ディープラーニングなら何でもできるんだと、真顔で言う人がいます。また、何でもいいから我が社で AI を開発せよ、という上司もいます。エンジニアの人から「私どもの上司がこんなことを言うんですよ」というような話しを今でも聞かされます。AI の性能は、知識系 AI では「形式的知識の質と量」、データ系 AI では「データの質と量」に依存するのであり、機械学習やディープラーニングが適用できない、ぐちゃぐちゃの悪構造データが圧倒的に多いので、その構造の整い具合（質）が問題になるのです。当然、データ量も問題になります。あと計算パワーです。さらに、ディープラーニングはかなり高い計算パワーがないと良い結果が学習できません。こういう事を理解した上で、ユーザの要件をどの程度実現できるか、それを考えていく必要があるわけです。今言

ったように機械学習やディープラーニングでは、データハンドリングにコストの6割がかかります。実際に機械学習、ディープラーニングを実行するには、ライブラリが充実してきて、新たなコーディングはあまり必要でないで、コストは1割程度で済みます。そして学習結果を読み解く必要があって、色々な変な結果も出てきますので、これは使えるけどこれは使えないとか選別しないとダメです。この評価をするのは、人間ですからそこにまたコストが3割かかります。単なるソフトウェアメトリクスではなくて、サービスの話がすごく関係し、総合評価する必要があるという事です。すなわち、データサイエンスでは、データ前処理：機械学習：結果後処理=6:1:3程度のコストで進むのが、現場の機械学習なのです。

AIという大括りで評価しないで、技術レベル、せめて先程お見せしたこのレベルまで勉強して、それぞれの適用可能性と限界と適用条件とユーザが提示した要件を比較し、ここは擦り合わせできるけど、ここは無理、というような判断をしていくことが必要です。そうすると変なAIプロジェクトにはなりません。メーカーの人に聞くと、AIプロジェクトの半分はつぶれてしまうと言われます。これはユーザの無理解、変な要求のためである、ということです。ユーザ側は、プログラミングが分かる必要はないのですが、マクロレベルでAI技術を理解する必要があります。

■行政オープンデータ+ディープラーニング→日本の改革案！？

ちょっと変わった応用例をご紹介します。行政オープンデータとディープラーニングを連携させて、日本の将来をどうすべきか改革案を出そうというテレビ番組がありました。こんな事は実際できるはずもないのですが、NHKのEテレ特集番組で、ユーザの実際の要求ではなく、仮想的な要求なのですが、2年ほど前に放送されました。行政が提供する700万のオープンデータにディープラーニングを適用した結果、日本の改革案として、「40代の一人暮らしが日本を滅ぼすのだ」という結果が出てきました。意味がわかりませんが、さらに「少子化を止めるには、結婚するのではなく、車を買った方が良い」とか、「病院を減らすと健康な人が増えるので、どんどん病院を減らそう」というような、意味不明の結果が学習されました。この時ゲストで出演していたマツコ・デラックスが「お前は何を言っているんだ。訳が分からない」と怒ってしまいました。NHK技研の技術者と少し付き合いがあるのですが、しっかりした技術をお持ちで、ディープラーニングの学習では因果関係が出てこないで、原因となる変数が隠れて潜在変数となっており、学習結果は、疑似相関に過ぎないなど、番組では説明されていました。例えば、最後の学習結果では、夕張市の行政データを使ったようですが、夕張市は破綻して多くの大病院が札幌などの大都市に転居した結果、夕張市の患者もそこに転居してしまった結果、夕張市の患者数はかなり減少して「病院が減ると健康な人の割合が増える」という結論を出したわけです。従いまして、学習結果を人間が因果レベルで読み解くことが必要にな

ってきます。読み解く AI も将来はできるかもしれませんが、現状では無理なので、人間が因果関係を読み解く必要があります。この点は、番組でも言及していて、ポイントを押さえているなぁと思いました。ただ「最後に妥当性は人が検証しないとイケない」と説明した後に「何かわかりやすい神のお告げでもあるわね」ともコメントし、そのコメントが独り歩きして、私の所に相談に来られた方が、「何か神のお告げと書いていましたよね。ディープラーニングってすごいんですね」と番組の感想を述べられ、誤解・混乱が起こっているなぁと思うわけです。

■アルファ碁

アルファ碁という囲碁 AI ですが、これはちょうど 3 年前のシンポジウムで説明しましたが、李世ドル（イ・セドル）がアルファ碁と対戦し、アルファ碁に 1 勝 4 敗で負けてしまいました。李世ドルは世界ランキング 2 位ですから、「いやあ、AI はもうトッププロに勝つようになりましたよね。ということは、我が社の業務でしたら、AI は何でもできるでしょう。囲碁のような難しいものでもできるのですから」。というようなコメントをする人が出てきました。

囲碁とか将棋とか、ボードゲームで次の一手を探すのが、それがビジネスの意志決定に近い、というような言い方をされて、ボードゲームとビジネスの類似性を考えられるわけです。でも、そういう話ではなくて、アルファ碁の場合は、アルファ碁同士、ソフトウェア同士を並列コンピュータにより 3000 万回対戦させて、棋譜データを作っ

たのです。その結果、勝ち負けがつくので、勝った方の対戦記録の棋譜、こういう石の置かれ方をして次にこういう手を打つと、理由はわからないけど勝っている、そういうデータを集めてきてディープラーニングにかけると、新しい勝ちパターン（定石）が誕生したのです。その新しい定石に従って、打ってくるから、人間のトッププロ棋士は意味が分からないのです。「何だ、この場面でこの手は？」となります。アルファ碁側は 3000 万回の対戦記録からのディープラーニングによる学習結果に沿って打っているだけなのです。ですから、ビジネスの意志決定に関する大規模整構造データがあれば、AI 化できる可能性はありますが、通常は、そのようなデータはあまり無いので使えないということになります。けれども見かけの類似性だけで「囲碁のできるのだから、我が社の業務も全部 AI でできるでしょう」というユーザがまだにいるのは問題だと感じています。

■グループ討論支援ロボット SOTA

それでは、残りの時間で AI ロボットについて説明します。私は、現在、国から支援を受けて、AI ロボットの研究を 5 年程度、続けており、今年度で終了予定です。

実践テーマは 2 つあり、システムの詳細は割愛しますが、先程、説明しました IBM のプロジェクトディベーターまでは到達していませんが、児童達のグループ討論に AI ロボットを参加させて、言葉と言葉の意味ネットワークであるオントロジーとナレッジグラフを使いながら、児童達の議論の中に入れていくのが一つの実践テーマです。

昨年「地球温暖化で僕たち私たちにできることは何？」というテーマで、杉並区浜田山小学校でこの12月に実施した実験について説明します。いろいろな機能があるのですが、ポイントは地球温暖化に関する専門用語の意味関係を書いたものを使って、いろいろ助言することです。それ以外に、時間や発言回数を記録し、発言回数が少ない子には、発言を促したりします。私も久しぶりに小学校の自由討論の様子を見ましたけれども、あるグループでは、一人の児童だけがよく発言し、他の児童は、それに押されて沈黙が続くケースがありました。この場合、あまり発言していない児童に「〇〇君、あまり発言していないけど、どうなの？」のようにロボット SOTA が声をかけます。そうすると話し続けている児童は「俺はちょっと喋りすぎたかなあ」と気づき「おい、お前、発言の番だよ」とか言って、チームワークが形成されていきました。発言回数を記録しているだけですが、チームワーク形成に貢献できたことは面白かったです。それでは、その様子を動画で見て頂きましょう。児童は、みんな首掛けマイクを1つずつ持って、SOTA は誰が喋ったかを認識できます。

<ビデオ再生>

生徒4, 5名 + SOTA 自由討論

今、二酸化炭素についていろいろ知りたいな、という状況になったので、この女子児童が「SOTA、二酸化炭素について教えて」と言ったら、二酸化炭素にリンクしている意味ネットワークの情報を提供して、説明

したわけです。マイクにより、A君が何を発言したとか、地球温暖化に関連するキーワードとして「二酸化炭素」「北極」「南極」など40語ぐらいをSOTAは記憶しており、児童がそのキーワードを発言したら、「〇〇君、今南極と言ったね、地球温暖化で南極が溶けると地球の海面は何と82cmも上昇するんだよ」のように議論を展開していきます。児童は、SOTAが自分に声をかけてきたことで「え〜、ほんと？ すごいな」というように好意的に知識を獲得していきます。ある意味、人とロボットの関係を構築する試みがなされているようにもとれます。この様子をご覧ください。

<ビデオ再生>

ちょっと声が聞こえづらかったですが、SOTAから声を掛けられた児童は、このようにガッツポーズをして喜ぶわけです。これは想定外でした。「〇〇君、今、△△って言ったね」というと、「わあ〜、え〜、お前どうして分かるの!？」というような感じで喜ぶわけです。そして、SOTAが「車に関してガソリン車以外に電気自動車などもあるよ」と発言した後、チームの議論は活発になっていきました。後半の議論にはロボットSOTAは参加せず、児童達だけの意見交換になりました。

「環境に優しい車って何だ」

「電気自動車だよ」

「でも火力発電所で電気を作ってそこでCO2出るから同じじゃないか」

「だったら太陽光みたいな自然エネルギーで発電したらいいじゃない」

「でも、まだ、それってできてないじゃん」といような感じで、このチームは議論が活性化していきました。

全体で 8 チームあり、最後に、すべてのチームが発表して先生が評価するのですが、このチームが一番優秀で、優秀賞をもらいました。そうしたら、このチームの児童達がロボット SOTA に近寄り「お前のおかげだ。みんなでお礼を言おう。ありがとう」と感謝してくれました。これはやや感動的な場面でした、このようなことは想定しておらず、逆に「お前の言っていることは、関係ないよ」みたいな感じで、SOTA は置いてきぼりにされるのではないかと心配していたのですが、児童と SOTA の間に良い関係ができたと言えます。

浜田山小学校で色々な実験をやっているのですが、全国に普及するために、今夏に「誰でも AI ロボット授業」というセミナーを開催し、小中学校の先生に教師ロボット連携授業の講習会を開催予定です。

■サービスロボット

2005 年、多くのサービスロボットが愛知万博で展示されたことから、今後、サービスロボットの時代になると期待されたのですが、実現されていません。万博向けのイベントで終わったわけです。2002 年に登場した床掃除ロボット Roomba (ルンバ) は世界で 2000 万台、2005 年に登場したアザラシ型のセラピーロボット PARO (パロ) 42 万円もしますが世界で 1000 万台売れています。パロは北欧で人気があり、デンマークの方が日本より早く商品化されたようです。42 万円もしますが、抱いているだけ

なのに心が癒やされるというメッセージが受け入れられたようです。そして 2014 年、Pepper (ペッパー) が販売され、銀行とか商店で業務案内をするペッパーを見かけるようになりました。このように、徐々にサービスロボットが普及しようとしています。人と対話しながら自然にインタラクションを技術には課題が多く、サービスロボット時代が到来したという程、普及はしていません。2015 年、日本はロボット新戦略を立て、サービスロボットの研究開発を進めています。自然なインタラクションの実現には色々と課題があるわけです。

■ペッパー

ペッパーはすでに 2 万台販売されましたが、半数以上は倉庫で眠っていると言われてます。販売最初は、物珍しさから客寄せ効果がありましたが、効果は数ヶ月で終わりました。その理由の一つが、処理が定型ルーチンなので、ペッパーの反応が分かっちゃって面白くないということがあります。ただ、定型処理であっても、ユーザがペッパーの使い道を徹底的に探らないために、ペッパーを倉庫で眠らせている側面もあります。

例えば、ペッパー利用の成功事例として、はま寿司さんの事例があります。はま寿司で 1 年半位かけてペッパーに接客を担当させてみました。最初、ペッパーに手を高く上げさせながら「いらっしやいませ」と対応させたら、モーターがすぐに壊れて止まってしまったそうです。当研究室でもペッパーを使っていますが、ペッパーの手は 100g 程度のものしか持てませんから、持ち

運びには使えません。一見、人の体のように見えますが、把持動作には全く使えません。実際に使ってみると大きな動作制約があることが分かってきます。はま寿司でもそのような経験を積み、大きな動作から小さな動作に変更し、手を少しだけ上げて「いらっしゃいませ」と対応させたら、長時間動作可能になったそうです。

このように、業務とペッパー機能の擦り合わせを何度も実行した結果、現在、500店舗でペッパーが座席案内で活躍するようになったそうです。はま寿司の受付付近の混雑が解消され、はま寿司の店員の業務効率は3割程度向上したそうです。小さな不具合に対しても粘り強く対応を続けて、達成できた成果と言えます。ただ、ペッパーはやはり故障します。故障時に、ペッパーが受付できないから、店長が店員に「ペッパーに代わって、座席案内をやって」と依頼したら、店員は「僕は座席案内の経験がないので、できません」と断られたそうです。未経験という理由から、人が座席案内できない問題は、人とAIの協働時代に起こりそうな課題ですね。人の教育の在り方も考えていく必要があります。

次は、ステーキハウスです。ステーキのお皿は1.5kgもあるそうです。腕がかなり疲れるので、運搬ロボットに配膳と下げ膳をさせようとしたのですが、配膳をさせた時は店員の移動を妨げることが判ってきました。そのため、白線を引いて、運搬ロボットが白線に沿って洗い場に下げ膳を行うことに特化したそうです。この結果、白線部分は運搬ロボットの通り道であると店員は認識して、店員は白線部分では注意しなが

ら移動するようになり、徐々に、店員と運搬ロボットの協働関係が構築されていったそうです。単純な使い方ではAIロボットとは言えませんが、業務とロボット機能のうまい擦り合わせの一例として捉えられ、ローテクでも、要件と技術を擦り合わせれば、十分サービスに貢献できるといえます。

■ロボット喫茶店

当研究室では、ここ数年、ロボット喫茶店を実践しています。昨年、大学祭で6種類のロボットを連携させたロボット喫茶店を2日間実践しました。その様子をちょっとご覧下さい。

<ビデオ再生>

ロボット喫茶店 in 矢上祭

2018年10月6日

これは、ロボットアームがクレープを作っているところです。先程の子供は少し緊張して、ペッパーとの対話がスムーズにできませんでしたが、数種類ある中から特定のソフトドリンクを何とか注文し、あと、クレープ生地に関連した注文もしました。これは浜松のベンチャーが作っているロボットマシン「Q」というもので、クレープの生地を10秒くらいでサッと作ることができます。「この生地の上に、似顔絵、ハート形、星形を描くことができますが、描いてほしいものがありますか？」というような注文をペッパーが取ります。ここではこの右側のお客さんが「クレープに私の似顔絵を描いて下さい」ということを注文したので、その注文データがロボットアームに送

られ、まず Q がクレープ生地を作って、ベルトコンベアのように隣の JACO2(ジャコ2)というロボットハンドの位置に移動させ、カメラでお客様の顔を撮って、輪郭線を抽出する、という処理を実行します。

クレープ生地生成には3分もかかるので、その間、ロボット喫茶店で働いている各種ロボットの業務内容と機能について、表情が変わるロボット Furhat (フルハット、スウェーデン製)が、口の形や表情を色々変えながら説明しました。

これが、3分後、JACO2 が描いた似顔絵です。どうでしょうか？鼻は少しつぶれています。まあ似ていますよね。絞ったチョコレートから「あなたの顔は55キロカロリーです」ともコメントしています。

この似顔絵作成機能は結構人気でした。Furhat は、日本ではここでしか見ることができないので、お客さんは、Furhat と一緒に写真を撮ったりしていました。

次に、これはトヨタのロボット HSR です。注文したソフトドリンクをテーブルまで運び、胴体と腕を伸ばして、お客さんに手渡します。HSR も人気でした。あとでアンケートをとると、人と直接触れあうロボットに人気があったようです。

最後に、これはゴミ回収ロボット SociBot (ソシボット) です。クレープ用のゴミ箱とソフトドリンク用のゴミ箱の2つのごみ箱を背中にかついで、喫茶店内を巡回しているだけですが、SociBot も表情が変わりますが、その顔が怖くて、子供達は「え～、何、この顔！」というような感じで近づかなかったのですが、慣れてくると、SociBot の後をずっと追いかける子供も出てきまし

た。

2日間、合計14時間、開店しました。来店者は253名でした。クレープは60名程度の人が注文しました。その結果、売上高17万円、利益8万円でした。焼きそばの屋台をやっている学生さんより稼いでいました。

あと、ロボット喫茶店に関連して、ウィーン大学と共同研究を実施しています。ウィーン大学では、ロボットレストランの青写真的なイラスト情報からビジネスプロセスに変換するような研究をしています。我々の研究では、今日はお話しできませんでしたが、PRINTEPS (プリンテプス) というユーザ指向の AI ツールを研究開発しており、ワークフローや判断ルールを日本語で記述すれば、Python コードに自動変換されます。このロボット喫茶店も PRINTEPS により開発しました。それで、ウィーン大学のツールと PRINTEPS を連携させて、色々なロボット喫茶店をイラストレベルでデザインすれば、様々なロボット喫茶店が勝手に動き始めるという目標を立てて、共同研究を進めています。

■AI と人・社会の関係

最後に AI と人・社会の関係について触れます。6年前にオクスフォード大学のオズボーン先生が Creativity (創造性)、Dexterity (手先の器用さ)、Social Intelligence (相手の気持ちを思いやれるか) という3種類の知能は AI にはできないだろう、と提言しました。野村総研はこの提言を日本の職業に当てはめ「AI ロボットに変わる可能性の高い100種類の職業」を報告しました。現在、これらの報告以外に、人の知能と AI を比較

し、AIに置換される職業に関する考察が多く出てきました。例えば、マッキンゼーレポートでは、(jobs, lost, McKinsey)で検索すれば、160 ページ程度のレポートを見ることができます。そこには2030年頃には全世界でAIロボットが8億人もの職業を奪うと書いてあります。このスライドは、横軸がGDP（対数スケール）縦軸がAIロボット置換率%です。何と日本の置換率が最も高くなっています。この1位は喜ぶべきか悲しむべきかですが、本来、人がする必要のない無駄な仕事が多い国が日本であり、AIロボットによってそれらの仕事が置き換わっていくと予想されているわけです。その理由も160ページのレポートに詳細に書かれています。我々はこの辺りの事ことを真剣に考えていく時代になってきたのかなと思います。

人とAIロボットが連携する方法は、国ごと、あるいは人ごとに違うだろうと思います。成功例として、このスライドは安倍総理大臣が工場見学されたものですが、作業者と双腕ロボットがうまく連携して、製品を組立えています。グローリーという釣銭機械を作っている会社で、15程度の生産ラインがあって、コストパフォーマンスを考えて、このラインだと作業員・ロボット連携により採算がとれるという判断で、2015年頃から実施されています。ただし、それ以外のラインは採算がとれないため、産業用ロボットだけのラインとか、複雑な作業が多いので人だけのラインになっているそうです。一つのラインだけ作業員ロボット連携が実現できたわけですが、要求と技術の擦り合わせを十分に実施した成果と言えます。

ます。

ロボットレストランをお見せしましたが、中国では2012年に世界で初めてロボットレストランが登場しました。7年前ですから、白線上を動いて運んでいますが、子供が「わあ、面白い」と集まってきました。ただ、ロボットが子供にぶつかりそうになったら、ロボットは止まるだろうと思っていたら、止まらず子供にぶつかってしまいました。ぶつかるし、運搬中にお茶もこぼすし、残念ながら、このロボットは半年で解雇されました。世界で初めて解雇されたロボットになりました。機能があまりにも不十分だったと言えます。なお、時間の関係で今日はお見せできませんでしたが、現在、中国国内ネットショップ1位のアリババと2位の京東(ジンドン)が、ロボットレストランで競争しています。

日本では、変なホテルという名で、ロボットホテル、ロボットレストランが展開され、従業員も解雇されています。しかし、ソフトクリームロボットは、プラスチックボードにより完全に隔離された空間で作っており、産業用ロボットの延長線上であり、人とのインタラクションは消滅され、機械化された印象があります。銀行もワトソンとかを入れて合理化を進めていますし、フコク生命では、既にワトソンを入れて34人程度解雇しています。このように、AI（ロボット）による人の置き換えが各分野で進み始めてきましたので、何かしらの形で社会にAIロボットが入ってくるのはもう確実ですが、職業を大雑把にとらえないで業務分析を細かく分割して、かつAI技術も細かく分類して「このビジネスプロセスはこ

の AI が使えるけれどもここまでが限界」とかいう感じで、人と AI の関係性を導いていけば協働関係が見えてくるのではないかと考えています。

以上で今日の講演を終わります。ありがとうございました。

(文責：編集委員会)