

# 人工知能、 ロボット、 人の心。

湯川  
鶴章



TheWave



人工知能研究に  
火がついた!  
これからは  
“人工知能”が  
大きく産業を動かす!

2015・冬

3ヶ月  
限定出版

旬本  
シリーズ

TheWave

人工知能、ロボット、人の心。

湯川鶴章

# 目次

## はじめに

### 第1章 人工知能の急速な進化が引き起こすビジネスチャンス

1. 人工知能が急に進化し始めた！
2. 50年来のブレークスルー
3. これから人工知能が仕事を奪う業種
4. 宇宙の謎は人工知能が解く
5. まずは産業的価値のある業種から変化
5. ロボットは進化するとウソをつくようになる？
7. 無意識の領域を再現できた
3. コストさえかければ自動翻訳は完璧になる
3. ドワンゴ人工知能研究所 山川宏所長
0. 脳を模したソフトと脳を模したハードの合体で、人工知能はさらに進化する
1. 人工知能の進化は人間の脳の進化？脳にチップを埋める必要なし
2. IoT x 人工知能で、人間の脳の限界を超える
3. クラウド（雲）からフォグ（霧）へ Fog Computingが開くIT新時代

### 第2章 協働型ロボットの今

1. ロボット台頭の背景に労働者不足
2. 飲食店向けロボットは実現するのか
3. ロボットはすべての仕事を奪うのか
4. 急展開する業界状況

### 第3章 ソフトバンクのPepper

1. 破格の価格でPepperを売る理由「目指すはクラウドAI」

2. [これだけ違うPepperとこれまでのロボット](#)
3. [変化に直面したときの典型的な企業の対応【携帯電話業界編】](#)

## [第4章 Watson](#)

1. [今後10年で人工知能は劇的に進化する データが爆発する領域を狙え](#)
2. [これまでにないものを発見する](#)

## [第5章 大阪大学石黒浩教授](#)

1. [人間に似ていなくても姿を補完する想像力が人にはある](#)
2. [人間に「心」はない？あるのは複雑なシステムの行動原理](#)

## [第6章 匿名仮想座談会](#)

1. [Deep Learningの価値とは](#)
2. [人工知能は人間の脳を再現できるのか](#)
3. [これからの仕事について](#)
4. [人間はロボットに征服されるのか](#)
5. [ロボットは自分で学習できるようになるのか](#)
5. [ロボットが奪う仕事、奪わない仕事](#)

## [おわりに](#)

1. [著者プロフィール](#)

# はじめに

情報の賞味期限は3ヶ月。この本は3ヶ月間限定の電子書籍である。

この電子書籍は、わたしが主宰している少人数勉強会「TheWave湯川塾」の講義録と、わたし自身の取材メモをベースに構成している。一応、簡単な編集は行っているが、素材をそのまま残したようなかたちになっている。

「とにかく旬な情報を」ということを主眼に電子書籍で緊急出版することにした。一番最後に追加した原稿は2015年1月14日に取材した際のメモで、出版のわずか3日前に書き上げ2日前に取材先のチェックが終わった原稿である。

ということで、1つの本として、内容、構成を吟味し、主張を一本にまとめたものではない。そういう意味で本としての完成度は低いと思う。そのことはご理解いただきたい。読み終わって「結局、この本は何が言いたかったのか分からなかった」というような内容になっているかもしれない。

言いたかったことは、ただ1つ。ロボットと人工知能の時代が、すぐそこまで来ているということ。それだけだ。あとは、読者お一人お一人が、この電子書籍の中に書かれた情報の中で、自分のこれからのビジネスや生き方に活かせる情報を1つでも見つけていただければと願っている。

「TheWave湯川塾」は、少人数で議論することがウリのテクノロジー系ビジネスマン向け勉強会である。少人数でも事業として成立するように、6回の講義で15万円という比較的高額な受講料設定となっている。この高額な受講料がフィルターになっているおかげで、非常に熱心な受講者が集まる傾向がある。そうした受講生による議論の場であるから、議論にも熱が入る。

この電子書籍のベースとなっている第24期は「人工知能、ロボット、人の心」をテーマに2014年11月4日から12月8日までの6週間に渡って開催した。第1回は、わたしが講義を担当し、それまでの取材で得た情報などをベースに、24期のポイントを解説した。

第2回から第5回までは、以下のような講師をお呼びして議論した。

11/10（月） 影木准子氏（カワダロボティクス）

11/18（火） 林要氏（ソフトバンクロボティクス）

11/26（水） 元木剛氏（IBM）

12/1（月） 石黒 浩氏（大阪大学）

第6回は、最終回で、みなで全体議論した。

第24期に参加した受講者は全部で5人。大手IT企業のトップエンジニア、ITコンサルタント、ベンチャー企業経営者、大手通信事業者新規事業担当者など。少人数だったこともあり、議論は白熱した。

この電子書籍では、その議論の内容も紹介したいと思う。ただ塾の講義の内容は、オフレコのものも多く、当然ながらオフレコ情報は、この電子書籍の中では紹介できない。

第4回のお阪大学・石黒浩教授の講義は、非常に刺激的な内容だったが、オフレコを条件にお話いただいたので、残念ながらこの電子書籍の中には掲載していない。代わりに、わたしが石黒教授を取材した際のメモを掲載させていただいた。オフレコ情報ほど刺激的な内容ではないが、石黒氏の主張の根幹をつかんでいただける内容になっていると思う。

またこの本の最後に「仮想匿名座談会」という章を設けた。取材が終わってからの雑談や、懇親会などの酒の席で出た本音のエッセンスをちりばめた完全フィクションである。だれの発言かは分からないし、特定の個人、企業、組織の非公開情報には一切結びつかない内容になっている。ただ人工知能やロボットの領域の最前線にいる研究者やビジネスパーソンがどのような考えを持っているのか、漠然と分かる内容になっている。ある意味、この本の中で最も価値あるコンテンツかもしれない。ぜひ最後までお読みいただきたいと思う。

それでは、人工知能やロボットに関する今最も旬でホットな情報をお楽しみいただければと思う。

2015年1月21日

湯川鶴章

# 第1章 人工知能の急速な進化が引き起こすビジネスチャンス

「50年来のブレークスルーが起こったんです」。ロボットと人工知能に関する僕の興味、関心は、松尾豊・東京大学准教授のこの一言がきっかけで一気に火がついた。

どうしたことなのかは、僕のブログの記事「人工知能が急に進化し始めた」に完結にまとめた。なので、まずは、このブログ記事をここに掲載したい。

## 人工知能が急に進化し始めた！

人工知能。何十年も前からある言葉だ。国家プロジェクトとして研究されていた時期もあった。それでも完成しなかった。やはり人間の脳は複雑で、それをコンピューターで真似することなど不可能かもしれない。

「ところがブレークスルーが起こったんです」と東京大学の松尾豊准教授は熱く語る。

### ▶2012年。人工知能研究に火がついた

2012年。人工知能の精度を競う国際的な大会で、カナダのトロント大学がぶっち切りの勝利を収めた。それも1つの大会だけではなく、3つ続けてだ。

「優勝したのは、画像認識、化合物の活性予測、音声認識など3つのコンペティション。まったく異なる領域にも関わらず、今までその分野を専門的に研究していた人たちを追い抜いてしまったんです」。Preferred Networksの岡野原大輔氏が解説してくれた。「それが立て続けに起こったので、コンピューターの各分野の研究者は、大きな衝撃を受けました」と言う。

トロント大学が開発したのは、ニューラルネットワークの分野の中の Deep Learning と呼ばれる手法。ニューラルネットワークとは、脳のニューロン（神経細胞）とシナプス（神経細胞結合）の回路を、コンピュータ上で再現したもの。人間の脳と同様に、正しい答えを出した回路が強化されるように設計されているので、コンピュータが自分自身で物事を学習していくことのできる仕組みだ。

そのニューラルネットワークを何層にも重ねるのが、Deep Learning と呼ばれる手法。「猫」という概念を理解するために一番下の層のニューラルネットワークが、直線や曲線を認識する。次の層で目や耳という部位が認識される。次の層では目や耳を含む顔が認識される。そして最後の層で身体全体が認識されて、「猫」という概念を理解する。そんな風な仕組みだ。

ドワンゴの人工知能研究所の山川宏所長は「人間がモノを見た場合、視覚から情報が入ってきてそれを脳内で階層的に処理していて、5層か6層のところで抽象的な表現が出てくるって言われています。人間の脳の中の処理のこの部分が、これまではコンピュータではまったくできなかった。それが Deep Learning でようやくできるようになった。長年、超えられない壁だったわけですから、すごくインパクトが大きい話です」と語る。

人間の赤ちゃんは2歳ぐらいで言葉を覚えるようになる。それまでの2年間で、物の概念をつかもうとしているのだという。家の中には、どうやら父親と母親という2人の大人が存在するらしいという概念を2年間かけて学習したあとに、「パパだよ」「ママだよ」と概念には記号があることを教わるので、初めて「パパ」「ママ」と話せるようになるのだという。

「この2歳までの脳の学習の仕組みをなかなかコンピュータで再現できなかったんですが、それが Deep Learning で可能になったんです」と松尾准教授はその意義を語っている。

## ▶ 2年間でさらに進化

「そして2012年から一気に、人間の脳を模した人工知能の研究に



火がついたんです。今、ものすごい勢いで研究が進んでいます」と岡野原氏は指摘する。

今年10月には、米Googleが、ニューラル・チューリング・マシンと呼ばれる技術に関する論文を発表している。

国際情報学研究所の市瀬准教授によると、記憶を統合できるようになった仕組みだというが、ドワンゴの人工知能研究所の山川宏所長によると、「まだソート（並べ替え）アルゴリズムくらいですが、プログラムを作れるようになるんです」と言う。山川所長は「自分でプログラムを書ける人工知能を作るとというのが、人工知能研究者の長年の夢だったんです。1970年代から80年代にかけては、そういう研究がいっぱいあったんです。ところがほとんどが挫折した。今回発表された論文によると、Googleのニューラル・チューリング・マシンでソートプログラムを書けるようになった。今までできなかったことができるようになったということで、意義は大きいと思います」と解説してくれた。

岡野原大輔氏によると、リカレント・ニューラル・ネットワークと呼ばれる技術も、2012年以降に大きく進歩した技術の1つ。静止画だけではなく映像や、テキストデータのような「系列データを扱えるようになった」のだという。

2012年にGoogleの人工知能が、インターネット上の1000万枚の写真を読み込むことで「猫」の概念を自分で学習したことが大きなニュースとなったが、今年秋にはGoogleの人工知能はリカレント・ニューラル・ネットワークの技術を使って、家族写真のような写真なら、何が写っているのか理解し、文章でキャプションをつけることができるようになった。わずか2年で、人工知能の画像認識能力と文書作成能力が一気に進化したわけだ。

▶「今後大きな山があるようには見えない」

松尾准教授によると、「Deep Learningで超えた山が、人工知能研究で一番大きな山。もちろん今後も課題はあるだろうが、このあとに大きな山があるようには見えない」と言う。

自分自身で学習できるようになった人工知能は、今後ロボットに搭載され、ロボットという身体を通じてさらに多くを学んでいくことだろう。また監視カメラや無人自動車、農業、流通、広告、医療、会計、介護、通訳、教育など、今後の人工知能の発達に合わせて、多くの業界が影響を受けていくと、松尾准教授は指摘する。

ロボット工学の権威、大阪大学の石黒浩教授は、人工知能やロボットが普及することで「今後、物理的な仕事はどんどんなくなる」と断言する。

その変化は指数関数的に早まっていくとみられる。集積回路の素子数が毎年、指数関数的に増加しているからだ。指数関数とは、最初はなだらかな傾きが突然、急な傾きに変化する関数だ。シリコンバレーの著名投資家Vinod Khosla氏は「この変化は竜巻のようなもの。最初は小さいかもしれないが、すぐに大きくなってあらゆる領域を飲み込んでしまう。人々は直近の動向を見てそれほどたいしたことではないと思うかもしれない。最初はとてもゆっくりとした変化だから」と語っている。

人工知能とロボットに仕事を奪われる時代。しかもその変化は指数関数的。変化に気づいた次の瞬間には竜巻に巻き込まれていた、ということになりかねない。

## ▶「IOT + 人工知能」の覇権争いはこれから

しかし変化は、チャンスでもある。Preferred Networksの岡野原氏は、これを大きなチャンスと見る。

「スマホ＋クラウド時代から、新しいパラダイムへの移行が始まります。そして1つの時代の覇者はその時代の成功に縛られて、次のパラダイムには移行できないケースが圧倒的に多い。今は、すごいチャンスのおかげです」と言う。

次のパラダイムとは何なのか。岡野原氏は「IoT（モノのインターネット）＋人工知能」が次のパラダイムだと見ているようだ。「あらゆるアプ

りやサービスは、バックエンドで人工知能につながるようになります」と指摘する。

パラダイムシフトは、業界の勢力図を大きく塗り替える。勝負の分かれ目は、今だ。

ブログ記事はここまで。どうでしたでしょうか。

この記事は非常に大きな反響を呼び、この電子書籍の原稿を書いている時点では、Facebookの「いいね！」が1268件、Twitterのツイートが995件、はてなブックマークが901件ついている。特に最初のころは、はてなブックマークで話題を集めた。はてなブックマークのユーザー層は、エンジニアや、技術に興味のある人が比較的多いと言われている。まずは、技術者に受けたようだ。その後はTwitter上で話題となり、最終的にFacebookユーザーにまで広がったようだ。

さてあらゆる記事は、取材で得た情報の氷山の一角を書き記しているに過ぎない。完結にまとめることで犠牲になった情報は山のようにある。このブログ記事には書かなかった情報が、この本のコンテンツとなる。

まずは松尾准教授とどのようなやりとりをしたのか。その取材メモをここに掲載したい。

取材日 2014年11月7日

場所 東京大学工学部2号館

取材対象 東京大学大学院工学系研究科 松尾豊准教

授

## 50年来のブレークスルー

——まずは人工知能の定義から教えていただけませんか？

確かにいろいろなレベルの定義があるかと思います。「人工知能」と

銘打っているもの、呼ばれるものには、技術的に見て幾つかのレベルがあるのではないかと思っています。

レベル1は、最も技術的に簡単なもの。今までの機械よりも、少し人間っぽいかなということで「人工知能」とメーカー側で呼んでいるものなんかがそうですね。「人工知能搭載エアコン」とか。まあ、これはマーケティング用語としての「人工知能」だと思います。

レベル2は、実際に人間の脳のような働きをするのですが、表現を人間が与えないといけないタイプ。「弱いAI」と呼ばれているタイプです。ただ1956年に人工知能の分野が出来て10年から20年の間に、そういう研究はだいたい終わっています。ただ最近はコンピューターが高速になり手法が多少改善されたということで、こうした「弱いAI」がかなり賢くなっています。ただ基本的なアルゴリズムは昔のものと同じです。IBMの「Watson」やスタンフォード大の医療診断マシンなども基本的には昔のアルゴリズムなんです。そこにウィキペディアなどの情報源を加えることで、すごく賢くなっているわけです。

レベル3は、表現学習のところの原理が大きく変わっています。「強いAI」と呼ばれるタイプの人工知能です。何を表現するのかというところをコンピューターが自動的に取ってくるんです。

## ——強いAI、弱いAIという表現は一般的？

はい、もともとはサールという人が提唱した概念です。強いAIには心が宿ってるって言いました。弱いAIは、チェスをするとか部分的な領域が得意なAI。もともとはそういう意味で使われていました。

## ——IBMのWatsonについては、どう思われますか？

Watsonはすごいと思います。マーケティング戦略としては素晴らしいと思う。ただ技術的には、クエスチョン、アンサリングという領域はずっと前からありました。iPhoneのSiriなんかもそうだけど、昔からある原理です。日の目を見ずに研究者が諦めていたところに商用化に成功して、よかったなって思っています。

クエスチョン、アンサリングは、精度をちょっと上げるのに、細かいルールをたくさん作らなければならなくて、しんどいんです。改良は可能だけど、疲れる領域だった。なのでIBMのWatsonのように、華々しくクイズのチャンピオンに勝ちましたと宣伝してくれると、研究者にとって元気

が出るので、ありがたいです。またWatsonは、回答の速さと精度を上げるために工夫をしている。考えてやってるんだなと思います。ちゃんとした技術の進歩が中であつたんだなと思う。純粹にすごいと思います。

——Deep Learningがすごいという話をよく耳にするのですが、  
いったいどんなものなんでしょうか？

50年来のブレイクスルーです。

50年来の難問は「表現を獲得する」ということ。Deep Learningが、それを解決する手がかりを与えてくれました。

——「表現を獲得する」とはどういう意味なんでしょうか。

これまでのコンピューティングはデータを扱うときにデータを前処理、加工して、変数に直すしかなかった。その前処理によって、そのあとの処理の精度が大幅に変わってくるんです。

人間の場合はその前処理を無意識にやっている。

視覚にしる、聴覚にしる、それ以外のデータにしても、人間の場合は、なんとなくここが重要だということを、しばらくすると自然につかんでしまうんです。

それを今までのコンピューターは、できなかった。ところがDeep Learningはそれを可能にするめどを示しています。

例えば画像の場合、画像の中から特徴量を自動的に切り出すことができる。特徴量とは、簡単なものだと、エッジの検出、角の検出、などのようなものから始まって、それが組み合わさることによって「顔がある」「人間の顔」「猫の顔」などという高次の特徴量までつながっていく。そうすることで高次の特徴量も自動的に取り出すことができるわけです。

——どういう仕組みで、高次の特徴量をつかんでいくのでしょうか？

オートエンコーダーというものを使います。

従来の機械学習は、教師あり学習、教師なし学習という分け方をすることが多いのですが、Deep Learningは教師なし学習だけれども、一見すると教師あり学習のような扱いをするんです。どういうことかと言うと、例えば画像のデータを与えた場合、コンピューターは与えられた画像のデータの一部を消して、「その消えた部分を残った部分から推測しなさい」という問題に変えて、自分自身に質問する。画像を1つ与えるだけで、たくさんの擬似的問題を作るわけです。それをニューラルネットワークで解かせていくと、ニューラルネットワークの隠れ層にあたる部分に特徴量が自動的に獲得されてくる。

それは画像の一部を見て、残りを当てるということにおいて重要な特徴量なので、画像を端的に表わす特徴量が自動的に選ばれる、というわけです。

## ——具体的な画像の例だとどうなるのでしょうか？

コンピューターに「1」という数字の絵を与えたとしましょう。コンピューターは「1」という数字の絵の下半分を隠し、「隠された部分を推測せよ」という問いを自分自身に出す。下半分が右に傾いていると推測する回路もあるだろうし、左に傾いていると推測する回路もあるでしょう。下半分が○になっていると推測する回路もある。

そうした解答の中から下半分をまっすぐな線と推測した回路が「正解」となります。同様に上から7割ぐらい隠れた絵を作ったり、4割ぐらい隠れた絵を作ったりして推測を続け、正解に共通した特徴をつかみ出す。「1」の場合だとまっすぐ下の伸びた線、ということになります。

その特徴を持つ数字を「1」とコンピューターに教えてあげれば、次から「1」を認識できるようになる。

## ——なるほど。このブレークスルーはいつ起こったのですか。

2006年にカナダのトロント大学のヒントン教授が論文で示しました。でもアイデア自体は古くからあったんです。日本では1980年代に福島先生という方が、ネオコグニトロンという同じようなモデルを示しています。同じようなモデルを考えた人はたくさんいました。

ただ2012年の機械学習の画像認識の国際的コンペティションで、ヒントン教授のチームが圧倒的な成績で優勝したんです。コンピューターによる画像認識の精度は、エラー率で表されるのだけど、それまでのエラー率は26%前後でした。ところがヒントン教授のチームはエラー率を16%にまで下げることに成功。2位のチームのエラー率26%に圧倒的な差をつけたんです。

Deep Learningのコンセプトが実際に画像認識の精度を大幅に向上させたわけです。

わたしは、ウェブ上のデータ、企業のビッグデータが専門でやってきたので、この領域でDeep Learningが使えないかを研究しています。それとDeep Learningはまだまだ欠点があったり、発展の余地があるので、時系列方向に発展させていこうとしています。

——過去がこうだから、次はこうなると予測するという話ですね。

そうです。

——Deep Learningと、今までのコンピューティングの違いは？

2つあります。機械学習という文脈で言うと、これまでは変数を人間が作らないといけなかったんです。前処理をして、いい特徴量を作ってあげると、機械学習ができる。

——コツを教えてあげないとだめだったという話ですね？

そうです。一方で記号処理の世界。推論、探索、という世界は、記号は扱うけど、記号がどういうデータに紐付いて現れてくるのか、記号と指すものの関係を捉えないままに、記号だけで処理をしてきたんです。

なので、フレーム問題だと、ロボットがどういうふうに動くか、というルールを与えたときに、例外的な事象が起これば困ってしまう。

シンボルがシンボルだけで成立していて、裏側がないんです。

——九官鳥が、意味を理解せずに言葉を喋っているように感じにはなりません。ですが、実際には意味を理解できていない。そんな感じでしょうかね。ところがDeep Learningでは記号同士の関係性についても分かるようになってきているわけですか？

いったん記号の概念が分かれば、あとは簡単なんです。

人間の赤ちゃんは、2歳ぐらいに言葉を覚え始めます。それで一気に言葉の数が増えていくんです。そこに至るまでに赤ちゃんの中で概念が獲得されているわけです。大人に教えられなくても、概念をつかんでいる。前処理をしているんです。

前処理の部分を、人工知能ではDeep Learningがやってくれるんです。

——じゃあ、今は2歳くらいまでの脳の動きを再現できるようになったということですか？では、これからまだまだイノベーションが必要なのでは？

いや、2歳以降というか、専門的知識の取り扱いの部分はもうできているんです。基本的には、これですべてできるんじゃないかって思っています。

——そうか、専門的知識の取り扱いって、今のコンピューターが得意とするところですね。これまで難しかったのは、前処理の部分。そこができるようになりそうなので、これで一気に進みそうだってわけですね。でも人間の脳ってまだまだ分からないことがあるんじゃないでしょうか？

人間がほかの哺乳類と違うのは、大脳新皮質のところなんです。大脳新皮質以外がやってることは、運動、感覚、記憶、行動計画などです。こういうのが大脳新皮質と連携することによって、知的な行動を生み出しています。つまり、その組み合わせなんです。



それは、Deep Learningと、例えば強化学習とかプランニングを組み合わせて、対応できるはずで、そこはこれから研究がされていく分野だと思う。すごくおもしろい領域ですね。

## ——人の心って、実はどういうものなのかは分かっていないのでは？

心が何なのかという問題がありますが、人工知能の研究的にはあまり関係ないかなって思っています。人工知能はあるタスクを知的に解ければいいので。

心がなにかというのを解明したいというのは、医学的興味じゃないでしょうか。工学的には作りたい、という興味。なので、ちょっと違う領域だと思います。

心を解き明かそうとすると、恐らく人工知能を作っていく中で、外界をシミュレーターとして作るという機能が必要になってくるはずで、その外界のモデルの中で、自分がどういう行動をすればいいか。行動計画を立てるときにシミュレーターが必要なんです。シミュレーターの中には自分がいて、自分の行動は将来の時間を含めて自分で決めることができる。自分以外のことは、こういうふうに動くと予測するしかない。その中で、自分というものの特異性が自分であり、心である。心はそういうものだという解釈ができるかもしれません。

心が何であるかが知りたいというのであれば、それに対応するような人工知能の必然的な機能が必要。それが現れてくると、それが心であると言えるようになるんじゃないかと思えます。

まあ人間の脳を知りたいという立場と、知能を作りたいという立場。立場が違ってるとだと思います。

知能を作りたいという立場からだと、予測するという機能が大事になってきます。予測能力を高めるために、そういう表現を作り出す、行動を計画する、というこが必要なんです。

人間を理解したいという立場からすると、本能だとか、人間独自の視覚、聴覚、嗅覚、味覚、触覚、そういうのを含めて理解しないといけない。そうするのは、ちょっと難しいというより、興味の方向が違ふんですよね。

——なるほど。飛行機を作りたいのか、鳥を作りたいのか、というような感じなんでしょうね。工学側の研究者は、なにも鳥を作りたい

と思っているわけじゃない。そういうことなんでしょうね。ところで脳って、電気信号と化学変化だけで本当に再現できるものなののでしょうか？ほかになにか要素があったりしませんでしょうか？

基本的に、電気信号と化学変化でシナプスが強化されると思います。そこに対していろんな人がそうじゃないことを信じたいので、量子現象があるとか、いろいろ言う人はいます。でも普通に科学的で合理的な人だったら、電子信号と化学変化でプログラムできると考えることが妥当なんじゃないかなと思います。

——そうですね。僕も人間の脳は電気信号と化学変化だけじゃないって信じたいほうの人間ですね（笑）。人間の心が、何か特別のエネルギーを発信しているというような発見があればいいのに、って願ってますから（笑）。でもまあ普通に科学的に考えれば、脳って電気信号で表現できる、ということなんでしょうね。では潜在意識はどこにあるんでしょうか？

明示的な意識からアクセスできないメモリ空間が、潜在意識ではないかと思っています。ニューラルネットワークの中でポインターが張られているところが顕在意識で、張られていないところが潜在意識。そこでも発火されているだろうけど、ポインターが張られていないので意識できないんだと思います。

僕は、潜在意識よりも、なぜ顕在意識が存在するのかのほうが興味がありますね。

言語化するときに、何にシンボルを与えるのかということと、顕在性、潜在性は関係があるのかもしれない、と思っています。なにもかもポインターをはって顕在的な情報を取れるようにすると、その数が多くなり過ぎて、言葉の数が増え過ぎる。言葉は、ほかの個体とコミュニケーションして合わせることで概念のロバスト性を担保しているので、たくさんあり過ぎると一致しづらくなり概念のロバスト性を担保できなくなる。そういうことで調整されているのではないか、と思う。割りと合理的に説明できるものだと思います。

これから人工知能が仕事を奪う業種

——なるほど。他の人とコミュニケーションする必要のない脳内の活動に関しては、明示的な意識からアクセスできないようにしておいて、他者とのコミュニケーションを円滑にさせる目的がある、という話ですね。おもしろいですね。さてこれからの松尾先生が考える技術的ロードマップはどのようなものになるのでしょうか？

次のようなロードマップで技術革新が進むんじゃないかって思っています。

(1) 画像から画像特徴の抽象化

(2) 音声などマルチモーダルなデータの抽象化

(3) 自分の行動のデータと観察データを含めた抽象化（赤ちゃんが0歳や1歳のときにやっていること。前に進む、ドアを開けるなどといったことができるようになる。自分が動くことで、モノが変化することを理解する。行動計画につながっていく）

(4) 行為を介しての抽象化（視覚、聴覚だけでなく、自分が触ったり動かしたりすることによって、カテゴリーをさらに分類する。ガラスを認識するには、ガラスは割れやすいという認識が必要。自分が動くことで、モノが変化するという事を利用して、モノを理解していくステージ。これによって壊れやすいものをやさしく触るなどができるようになる）

(5) 概念を抽象化できるので、それを言語にマッピング（モノの存在を理解できるようになったので、そこに言葉を紐付けていく）

(6) 言語理解した上で、ウェブや本を理解し、ますます賢くなる（言葉とモノが紐付けられたので、より多くの知識を得ていく）

——なるほど。本当に人間の子供が学習していくような順番で人工知能も進化していくわけですね。非常におもしろいですね。ところでこのロードマップは、人工知能の研究者の間での共通認識となっているんですか？

いや、そんなことはないです。あくまでも個人的な見解です。でも、こんな感じで実際に進んでいますね。今は3番目くらいまでできています。

——では、このロードマップに従って、いろいろな職業や業界に影

## 響を与えていく可能性がありますよね。具体的にどのような業種に影響が出てくるのでしょうか？

まずは画像の認識精度がよくなるので、広告や視覚による診断なんかにインパクトがあるかもしれませんね。

次に視覚だけでなくいろんなモードに人工知能が応用できるようになると、防犯、感情の領域にも人工知能は入っていくでしょうし、企業の購買データなどビッグデータから起こっていることを理解できるようになるかもしれません。

3番の段階になると、行動ができるようになります。農業や物流、ロボットが関係してくる領域ですね。

4番は行動を介した抽象化なので、やさしく触る、壊れやすいものに注意するということができるようになります。日常生活の中に入ってきやすくなるわけです。家事、介護、他者理解、感情労働などの領域に影響を与えそうです。

今までのコンピューターって例外に対応するのが得意じゃなかった。でも今、進化し始めた人間の脳を模した人工知能だと、自分で認識して判断できるようになるわけです。なので例外への対処も可能になります。つまり今までのコンピューティングとDeep Learningでは、環境の変化に対する対応の速さが違うんです。例えば異常気象で気象条件が変わった、道路交通法が変わって運転のルールが変わった、山岳地帯で通過できる道路が変わった、などという例外的な状況でも、人工知能は環境データから重要な点を抽出し自分でルールを即座に再構築します。

今のコンピューティング技術だと人間が最初からルールを作りなおさなければならぬので、急な変化に即座に対応できないんです。

翻訳って今でもある程度できたりはしますが、やはり完璧じゃない。例えば日本語と英語で同じことでも表現の仕方が違うことがある。そういう場合でも意味が分かっているので、どういうことか理解した上で翻訳できます。

——なるほど。なんとなく分かります。「girl」って「少女」って訳されることが多いけど、英語の「girl」って大人の女性が含まれることもあるんですよね。「昨日、バーで、とてもセクシーなgirlに会った」っていい方を英語ではするわけです。この場合のgirlは「少女」じゃないということを、人間ならすぐに理解しますが、「girl=少女」と記号で覚えているコンピューターは、間違っって訳してしまいますからね。

はい。翻訳って行間を読んだりしないといけないので、実は非常に難しいんです。秘書業務もそう。言われたことだけやっているようではだめで、あらゆる状況をすべて把握した上で最適な行動を選ばないといけない。人工知能にこれができるようになるのは、2030年くらいでしょうか。

**——これから直面しそうな技術革新の壁ってあるのではないのでしょうか？**

技術的に見ると、今までの人工知能の発展が、表現学習というところで抑えられてきたんです。表現を獲得するということができないから、いろいろやってきました。ところがそれを越えることができたので、技術的に見ると、この先に大きな山があるように見えません。社会的にそれを許容するべきかどうかという議論は今後あるかもしれませんが。技術的には、ここが難しそうだなというのはそれほどないです。当然、それぞれ難しいんですけど、やればそのうちできるだろう、という感じなんです。

**——シンギュラリティ（技術的特異点）についてはどう思いますか？人工知能の能力が全人類の能力の総和を超えるという時点があと30年ほどでやってくるって言われます。そのシンギュラリティの時点を超えれば、人間は人工知能に征服されるのでしょうか？**

シンギュラリティを2045年と予測する人も、2028年と予測する人もいますね。理論的には、AIが自分よりも賢いAIを作り出せるようになったときに無限に発達してしまう、ということは考えうると思います。ただそこにいたるまでにいろんなことが起こる。例えば予測精度が上がるので、金融市場に大きな影響を与えるでしょうね。AIを使った企業の生産性が上がるということもあるでしょう。仕事がなくなる人が出てくる。所得の再分配をどう考えるのか、ということも大きな問題になると思います。

シンギュラリティ以前に、こうした問題を解決しなければなりません。解決しようとする中で、「シンギュラリティまで行かさない」という決断を人類がするかもしれません。

シンギュラリティを超えてもいいのかという議論が人工知能学会の中

でもあります。考え方によって宇宙派、地球派と呼ばれて区別されることがあります。

地球派の考え方は、この世界は人間が主役で人工知能は人間のためにあるべき、というもの。宇宙派は、人間が人工知能を生み出してしまったら、人工知能の世界になる。それは止められない。人間は人間で幸せに暮らしている、というような考え方。

——松尾先生は、地球派、宇宙派のどちらなんですか？

さあ、どっちなんでしょうね（笑）。

——人類にとって非常に大きな変化が目前に迫っているように思うのですが、これに対して社会はどう対応すればいいのでしょうか？

人工知能に関わらず、IT社会でいろいろな新しいルール、決め事が必要になってきています。例えばネット企業はどこで税金を払うべきか。ネットの情報に対する「忘れられる権利」はあるのか。高齢化する中で投票権が一人一票のままでいいのか。

いろいろなところで新しい社会システムが必要となってきたんです。もう一度見直そうという機運になればいいんですが、果たして日本でそうなるのかどうか。

でも社会全体として、どう対応していくのが大事ということは間違いないでしょうね。

## 宇宙の謎は人工知能が解く

——ところで宇宙の謎は、人工知能が解くようになるのではないかって思っているんですが、どう思われます？人間には認知能力に限界があって、現在、過去、未来が同時に発生しているという概念を、直感的には理解できないですし、5次元の世界のことも想像できない。電波も見えないし、人間には見えないエネルギーがた

**くさんあります。人工知能だと人間特有の認知能力の制限を受けることもないので、自由な発想で宇宙の謎に挑戦できるのではないかって思っているのですが。**

まったく同意見です。僕は昔から数学、物理が好きだったのですが、物理学はオブジェクトを認識して、その関係を方程式で記述する。そのオブジェクトの認識が、人間の認識に依存しているんですよ。そうではなく、波のようなものを理解する、モアレ（干渉縞）のようなものをそのまま理解する、ということがあっていいはず。そういうような認識能力があったときに、この世界はどのように見えるのだろう。すごく興味があるところです。もし仮に人工知能がそういうように世界を認識して、われわれに認識できない世界の謎を問いてくれるとすれば、知りたいし、見てみたいと思います。

**——人工知能が発達すると、宇宙の理解がすごく速く進むのでは？**

実は友人の人工知能の研究者と話をしていて、どうして人工知能の研究を始めたのかという話になったんです。僕は、「人間ってなんだろう」という興味から人工知能の研究を始めたと言ったら、友人の場合は、もともと物理を専攻していて、物理学の進展の速度を見てると、自分が生きている間に宇宙を理解できるとは思えなかったそうなんです。それだったら人工知能を作って、それに宇宙のことを考えてもらったほうが早いのではないかと考えるようになったんだとか。シンギュラリティを先に起こして、宇宙論を作ったほうがいい、という意見でした。

**——おもしろいですね。ところで遺伝の仕組みは人工知能に必要なのでしょうか？**

そうですね。人工知能も死ぬ仕組みにしたほうがいいのではないかと考えています。世界を認識するのに、Deep Learningのように下のレイヤーから固定させていく方法だと、環境が変わるとその認識を変えたほうがいいのに、なかなか変わらないことがあるかもしれない。下から上に組み上がっているので、下のレイヤーを修正して上を変え

るのがすごい大変になってくる。それなら、いっそのことすべて壊してしま  
って、一番下から作りなおしたほうが早いんじゃないかということになり  
ます。コンピューターはいったん死んだほうがいい、という考え方ができ  
るわけです。

**——遺伝の仕組みを取り入れたことで、ウソをつくコンピューターが  
出てきたというスイスの実験がありました。そうなる可能性もある  
ということでしょうか？**

自己保存のプログラムを入れると、ウソをつくことで自己保存ができ  
る確率が高い状況の中ではそうなるかもしれないですね。

**——個体が死んだほうが種の進化という意味においてはいいとい  
うのは、自然界と同じですね。**

遺伝子の末端のテロメアという寿命の仕組みがあるから死ぬ、と言  
われています。その仕組みがなければ生きていられるんだろうけど、死ん  
で次の世代になったほうが種としての環境の適応能力が高くなりま  
す。種にとっては、そのほうがいいわけですね。

**——量子コンピューターが登場することで、人工知能の研究にどう  
影響するのでしょうか？**

僕はあまり関係ないかなって思っています。人によって意見は違うと  
思いますが、少なくとも人間の知能というものは通常の物理学の範  
囲内で収まっていると思うので、量子コンピューターを出さなくても、人  
間の知能、もしくはそれを超えるものは生み出せるんじゃないかなって  
思います。

**——人工知能の進化のロードマップって非常に興味深いのです**



**が、このロードマップに従って、人工知能が人間の仕事を奪っていくということになるのでしょうか？**

人工知能の進化で消滅する仕事、というようなテーマの本を書けば売れるんじゃないかって、冗談で言ってます（笑）。

人工知能はいろいろな業種に影響を与えるでしょうが、同じ業種でも、なくなる仕事、なくなる仕事があるんだと思います。

**——人間はどのような仕事についていけば安泰なんですか。究極の未来には、仕事はなくなるのでしょうか？**

そうですね。僕の予想では定型的な業務はどんどんなくなっていくんだと思います。会計、法律に関する仕事の一部はなくなっていきます。

でもクリエイティブ、大局的な判断が必要なところはまだ残ると思います。しかし、そうした仕事もそのうち人工知能がやるようになると、最後に残るのは人間と接する仕事じゃないでしょうか。例えばマッサージ師とか。

**——そこもロボットが取って代わる可能性があるのではないのでしょうか？**

ロボットがいいか、人間がいいか。ロボットがいいとなれば、確かにそれもなくなくなるかもしれないですね。

**——僕は人間のマッサージ師のほうがいいなあ（笑）。僕みたいなのがいる限り、人間のマッサージ師は安泰かもしれませんね。**

今後ロボットに関する倫理的な議論がされるようになるのだと思います。「人間にそっくりなロボットを作ってはいけない」だとか「人工知能も一日8時間寝ないといけない」とか。

——なるほど。そうすることで、人間の仕事を守ろうという動きが出てくるかもしれませんね。でももし最後にすべての仕事が無くなったときに、人間は何をすればいいのでしょうか？

何もしなくなり、ごろごろする。南の島で飲み食いしているかもしれないし、生物は争うようにできているので、昔の貴族が蹴マリをしたように、未来もそんなことしているかも。そんな中でもそこに生きがいを感じることをするほうが、生存確率が上がると思います。まあいろんなシナリオが考えられると思います。

——学会でそういう話は出たりしますか？

今年初めてそういうセッションが設けられました。

——本格的な議論はこれからということですね。そういう時代になる前にテロリストが人工知能を利用するという脅威があるのではないのでしょうか？

僕自身は、シンギュラリティ以降のことを考えるよりも、それ以前のそういうことを心配したほうが良いと思っています。

技術的なイノベーションを少数の人が手にしたときに、いろんなことができてしまう。それに対してどういう手を打たないといけないのか。ある意味で原子力みたいなものなので、みんなで管理しないとダメだと思います。社会的に合意を形成していかないとダメ。

まずは産業的価値のある業種から変化

——広告キャッチコピーが人工知能に取って代わられるのは分かります。でも星新一のショートショートのようなものを人工知能が

## 作れるようになるのでしょうか？

人間的な世界の理解にどこまで近づくか、という話だと思うんですが、人間的な世界を理解するためには人間的な感覚（五感）、人間的な本能を持っていないといけなんでしょうね。そうすることで初めて人間に近い概念をコンピューターが持てるようになる。そうすると、人間が共感できるような文章をコンピューターが書けるようになるかもしれないです。

ただ僕は人工知能がそういうようなことをするようになるとはあまり思っていない。というのはロボットや人工知能を、人間に似せることで実現する産業的な価値もあるとは思いますが、でもそれよりも、人間で似せなくてもいいので高い産業的価値を出したいというケースのほうが多いと思うんです。似せなくていいので産業的価値を出したいというケースが先にくるんだと思います。星新一のショートショートを書くというのは、人工知能にとって結構ハードルが高い。でするので人間が感動するような本を書くという仕事は、比較的残りやすいんじゃないかと思っています。

——なるほど。人間に似せることが技術的に難しく、しかもそれができるようになっても産業としての価値があまりない仕事。つまり儲からない仕事の領域に対しては、人工知能やロボットを進化させようというインセンティブが働かない。なのでそうした領域の技術開発は後回しになってしまうだろう、ということですね。ということは、儲からない仕事ばかりが人間に残ってくるということになりませんか？

そういう言い方をすると、ちょっと悲しいですが、まあそういうことですね。

——でも別の言い方をすれば、お金になるかもしれないけど、つまらない仕事をロボットや人工知能が受け持ってくれて、人間はもっとクリエイティブで、やりがいのある仕事に専念できる、ということにもなりますよね。儲かるかどうか、収益が上がるかどうかは、マーケットの需要と供給の関係で決まりますから、ロボットが作り出すような製品の価格は下がり、人間が作る製品の価格は上がるようになるかもしれないですね。ロボットや人工知能が作りそうもない領域の製品で、自分にしか作れないもの、世の中の人から強く求

**められるものを作れば、未来でも十分に儲かりそうですね。ところで、本能のようなものは遺伝の仕組みを取り込んで、何世代も経験していかないと身につかないものなんでしょうか？**

そう思います。本能は進化の結果、組み込まれたものだと思うので、「きれい」「怖い」は相当作りこまれた感情だと思うし、進化的にいろいろなバージョンを用意して生存確率の高いものが残るというやり方じゃないと作り込めないと思いますね。

**——じゃあ、本能は人間がロボットに埋め込むしかない？**

埋め込んでも、大雑把にしか埋め込めないと思います。

**——なるほど。ところで、松尾先生の今の関心事は？**

そうですね。オールジャパンで、世界と戦いたいですね。ただ機械学習の日本人の研究者の層が薄いのが悩みです。米国のIT企業は、機械学習の効率が1%アップするだけでも大きく儲かる産業構造を作っているので、そのインセンティブが相当高い。効率をちよつとでもアップさせるために、がんばろうという土壌になっている。

一方で日本の場合、そこまで精度に敏感な業種ってあまりない。そういう意味で投資がしづらい傾向にあります。そういう短期的な投資効果は期待できないので、日本は国として長期的な投資をすべきだと思ってます。長期的には投資対効果が高い領域だと思います。

**——海外勢は、どの程度のところを進んでいるんでしょうか？  
Facebookの人工知能は、変数を与えてあげるタイプですか？  
自分で概念をつかむタイプですか？**

今は与えるタイプですが、自分でつかんでくるタイプのものをどんどん入れてこようとしていますね。

**——Deep Learningは日本がリードしていますか？**

いや、そんなことはないですね。

——国別ではどこが強いのでしょうか？

やはりカナダ、アメリカですね。

——でも日本は研究者の層が厚いので期待できるという話を聞いたことがあるのですが？

第5世代コンピューターの研究とかしていたので、潜在的に先を想定している人が確かに多いと思います。

以上が松尾豊准教授とのインタビューだ。松尾准教授との話の中で、遺伝機能を持つロボットの話が出てきたが、それは次のような話だ。過去に書いた記事を掲載しよう。

## ロボットは進化するとウソをつくようになる？

2009年にスイス連邦工科大学ローザンヌ校で行われた実験で、1000個のロボットを100グループに分け、特定のルールに基づいて得点を競い合うようにプログラムしたところ、ロボットは最終的には他のロボットにウソをつくようになることが分かった。

実験に使われたのは、お掃除ロボット「ルンバ」のような円盤型ロボットで、それぞれがセンサー、カメラ、青色ライトを搭載。また「ゲノム（遺伝子の集合体）」として、グループごとに異なる264ビットのバイナリーコードが組み込まれている。

センサーや車輪、青色ライトには、脳で言うところのニューロン（神経細胞）の役割をする部品が搭載されており、センサーには11個のニューロンと、車輪に2個、青色ライトには1個のニューロンが搭載されている。これらのニューロンをつなぐ33の接続がシナプス（細胞間の信号伝達部位）の役割を果たすようになっており、シナプスの接続の強さは、8ビットの「1つの遺伝子」で制御されている。そして264ビットの「ゲノム」がセンサーから得た情報をもとに、どう動くかを定める仕組みになっている。

実験会場には「良い箱」と「悪い箱」の2つが設置され、ロボットは

センサーで「良い箱」を探して、会場内を動きまわるようにプログラムされている。「良い箱」には発色ライトが搭載されているが、ロボットのセンサーは下向きに設置してあるので、「良い箱」に近寄らないと発色ライトを認識できないようになっている。そして何分かゲームが終了。「良い箱」の近くに存在するロボットから順に高い点数が付与されるルールになっている。

そして1000個のロボットのうち、得点数の高い200個のロボットの遺伝子コードを掛け合して新しいプログラムを作成し、それを1000個のロボットに再び搭載するという実験だ。

自然界で言えば「良い箱」は食料ということになり、食料の獲得に適した遺伝子が次世代に受け継がれるという自然界の仕組みを、ロボットで再現した形だ。

またロボットには遠くを見渡せるカメラと、青色のライトが搭載されており、第1世代のロボットはランダムに青色ライトを点灯させていたのだが、9世代ごろには「良い箱」を見つけると青色ライトを点灯させるロボットが生き残るようになった。「良い箱」の存在を互いに知らせることで生存率が上がったためだ。

ただ次世代に持ち越せるのは上位200個のロボットの遺伝子だけ。9世代以降は反対に、「良い箱」を見つけても青色ライトを点灯しないものが増えてきたという。

食料を見つけても他のロボットに教えないほうが生存率が高いのだろうか。500世代以降になると、ロボットの60%は「良い箱」を見つけても、青色ライトを点灯させず「良い箱」を独り占めしようという行為に出たという。

また10%のロボットは、「良い箱」を見つけてもいないのに青色ライトを点灯させて、他のロボットを惑わそうとした。また3分の1のロボットは青色ライトを信用して近づき、3分の1のロボットはウソと見抜き、青色ライトから離れようとするようになったという。

なかなかおもしろい実験結果だけれど、「ロボットは進化するとウソをつくようになる」という短絡的な話ではないと思う。最初に1000個のロボットのうち200個しか生き残れないように設定してあるので、競争が生まれ、その結果、ウソをつくことが有利になっているだけだと思う。食料を得たロボットの子孫が制限なく繁栄するように設定すれば、協力し合って食料を獲得したのだろうと思う。

資源に限界があるのなら、だまし合いも発生するけれど、技術革新で資源が無尽蔵に増えていくのであれば、協力するムードのほうが高まるのではないだろうか。

さてDeep Learningの話をもう少し詳しく知りたいと思い、人工知能に詳しいPreferred Networksの岡野原大輔氏にも話を聞いてみた。

材日：2014年11月27日

所：株式会社Preferred Networks（東京都）

材対象：岡野原大輔氏（同社副社長）

## 無意識の領域を再現できた

**——人工知能の研究に大きなブレイクスルーがあったと聞いたのですが。**

確かに大きな山は超えたと思います。人工知能の研究って1950年くらいから行われてきたのですが、一番最初に注目された手法は、ものごとを記号化し、記号を元に解きましょう、というものでした。というのはコンピューターって数式の処理が得意なわけなので、言語を始め、画像なんかも記号として表現すれば、その記号を元に数式として解けるのではないかと思われたんです。そしてその方法でいろいろと試みが繰り返されてきた。ところがどうもうまく解けない。

どうやら人間の脳は、かなりの部分の作業を無意識のレベルで処理しているようなんです。無意識なのでそれがどのような手順の作業かよく分からない。よく分からないので、コンピューターで再現しづらい。このため人工知能がなかなか発展できなかったんです。それがこれまでの流れでした。

人間の無意識の作業ってどんなものがあるかというと、言語なんかも無意識に処理していますし、文章を読んだりするのも実は無意識に処理している作業を含んでいます。

その辺が簡単そうに思っていたんだけど、実際にやってみると簡単なルールだけでは解決できないことが分かったんです。例外的な処理が多過ぎて、うまく処理できないんです。

画像認識もそうです。「犬」と一言と言っても、犬にはいろいろな種類がある。同じ犬種でも、毛並みや色でいろいろなバリエーションがある。人間はそれでも、犬を見かけると、それが犬であると一瞬で認識してしまう。で、どうしてそれが犬だと分かったんですかと聞くと「うーん、どうしてだろ。だってどう見ても犬だし」という感じで、うまく説明できな

いんです。人間は無意識の部分で、犬の認識作業をやってしまっているの、その作業工程を自覚できないんです。

——いちいち顕在意識で、この動物はこうこうこういう特徴があるから犬である、とは考えていないということですね。その辺は無意識レベルで一瞬にして「犬」と認識していて、犬を見た瞬間に「これは犬である」って分かってる、ということですね。そう考えると確かに人間の脳ってすごいですね。

そうです。それでその無意識レベルの作業って、人間の意識上に上がってこないから、どのような手順で作業が進められているのかよく分からなかった。

1990年代ごろから、自然言語処理の分野で機械学習ができるようになったんです。機械学習って、コンピューターに直接ルールや知識を教えなくても、たくさんのテキストを与えると、そのテキストから文法とかのルールを自分で見つけ出すことができるってことなんです。自然言語処理の分野で、機械学習である程度の文章の解読はうまくいったんです。

ただそうはいうものの、人間ほど文章の内容をしっかり理解しているかという、そうでもなかった。突っ込んだ質問をしたら、ちんぷんかんぷんな答えが返ってくることもあった。テキストだけじゃなく、画像なんかもそうでした。

——スマートフォンの音声認識なんかでも「今日の天気は」「今日の予定は」などと聞くと、割りと正確に答えてくれますが、「午後からの雲行きは」とちょっと違った言い方で、聞くと、もう答えられない。それが現状ですよ。

ところがそこにブレークスルーが起こったんです。ニューラルネットワークの分野で大きなブレークスルーが起こったんです。

——ニューラルネットワークって、ニューロンやシナプスといった人間の脳の回路をコンピューターで真似しようという試みですよ。



はい、ニューラルネットワークの研究は1950年や1960年といった昔からずっと続けられてきました。でもなかなかうまくいかなかった。ところがカナダのトロント大学のヒントン教授のグループが、大量のデータを使うことでニューラルネットワークの精度が格段に向上するという実験結果の論文を発表したんです。

**——人工知能研究の50年来のブレイクスルーだったんですね。それはどんなブレイクスルーだったんですか？**

それまでは「猫」の写真を見て「猫」と認識させましょうというやり方だったんです。ヒントン教授のやり方は、それを何階層にも分けて、少しずつ変更しながら最終的に「猫」という分類に到達するという方法です。

**——階層による認識というのは、例えば一番下の階層では、尖った耳やヒゲといった顔の部分を認識させ、次の階層ではそれらの部分を合わせたものを顔と認識し、最後の階層では顔や胴体を含めた全体を認識するというように、幾つかの階層に分けて、認識を積み上げていく方法のことですか？**

そうです。そのレイヤーの数が今まで1層か2層程度だった。何層もの深いレイヤーになると、途端にパラメーターをどうやって決めればいいのか分からない。非常に難しい問題だったんです。それがヒントン教授は画期的な手法で、その問題を解決してしまっただけでなく、何層もの深い層に分けて学習するので、「深層学習（Deep Learning）」と呼ばれるわけです。

**——なるほど。幾つもの階層に分けて理解する。階層を深めて学習する手法なんでDeep Learningって呼ばれるわけですね。**

人間の脳も、きれいに区分はされていないんですが、やはり5層や10層ぐらいの構造になって認識するようになってるんじゃないかって言われているみたいです。2006年にこの方法がうまくいったというヒン

トン教授の論文が出たわけですが、ニューラルネットワークで成果がなかなか出ない時代が続いたんで、当時は研究者もほとんどいなかったんです。

ところが2012年に入って、ヒントン教授のグループが、同じDeep Learningの手法を使って、画像認識、化合物の活性予測、音声認識といった3つぐらいの人工知能のコンペティションで、それぞれ優勝したんです。まったく異なる領域にも関わらず、今までその分野を専門に研究していた人たちを追い抜いてしまったんです。それが立て続けに起こったので、各分野の研究者が衝撃を受けました。そして、一気にそこからニューラルネットワークの研究に火がついたんです。今、ものすごい勢いでDeep Learningの研究が進んでいます。Deep Learningが起爆剤になって、一気に人工知能の進化が始まってると感じます。

## コストさえかければ自動翻訳は完璧になる

**——2012年のトロント大の世界大会での優勝が起爆剤になって、人工知能の研究にどんな進化が始まってますか？**

幾つかあるんですが、例えば人工知能がプログラムを読んで、そのプログラムが何を出力するかを予測するようなニューラルネットワークが出てきました。プログラムをまったく教えないで、たくさんのプログラムの文字の羅列だけを見せるんです。そうすると人工知能は、文字の羅列のどの部分が数字で、どこの部分が条件で、どこの部分が繰り返しのかといった構造を自動的に見つけ出し、そのプログラムの出力結果を出すことができるようになりました。ニューラルチューリングマシンと呼ばれる技術ですが、今年の夏に実現しました。

**——なるほど、人工知能がプログラミングを理解するようになったということですね。となると人工知能がプログラミングできるようになる。確かにこれってすごいことですね。プログラミングが人間の仕事ではなくなる可能性がありますね。**

またもう1つすごいイノベーションがあります。それがリカレント・ニュー

ラル・ネットワークです。

これまでのニューラルネットは、階層の下から上へ一方向に上がっていくわけです。いわゆるフィードフォワード型のニューラルネットなんですが、リカレントニューラルネットワークだと、ループを持っていて、ある程度長い系列を扱うことができるんです。

画像っていわば一瞬のデータですが、われわれの脳は一瞬のデータだけを認識するのではなく、時間方向で長い系列の可変長のデータを受け取って、それを認識する。可変長のデータをうまく扱えるようになったのが、リカレントニューラルネットワークという技術なんです。

**——例えて言えば、今までのニューラルネットワークだと静止画しか扱えなかったのが、リカレント・ニューラル・ネットワークで動画も扱えるようになった。そういうような感じですか？**

そうです。時系列データを認識できるようになったんです。今までは、時系列を始めとする系列データをうまく認識できなかったんです。例えばテキストデータもそうです。

人間の脳はテキストの文を最初から読んで、意味を認識できるわけです。「山田花子さんが八百屋へ大根を買いに行きました」というような文章を、人間は簡単に理解します。でもコンピューターの中で、こうした文章を表現する方法がなかったんです。ところがLSTM（Long Short Term Memory）のようなリカレントニューラルネットワークを使うと、文章の表現をコンピューターの中でうまく扱えるようになったんです。

その表現の方法は、「分散表現（Distributed Representation）」と呼ばれています。この分散表現という概念自体はヒントン教授が1969年代くらいから言ってる仮説なんですけど、ヒントン教授は、人の潜在意識の中で知識はいくつもの要素に分かれて表現されているのではないか、と言うわけです。例えば、猫はAという要素が0.3、Bという要素が0.5、Cという要素が0.7などといった具合に幾つかの要素が合わさって、猫というものになっていて、犬はAという要素が0.4で、Bが0.6、Cが0.2というように組み合わせられて犬になっている。そんな風な考え方です。

**——例えば、Aという要素は「四足で歩く」という概念で、犬も猫も**

**この要素のポイントは高いが、猿は2本足で歩くこともあるのでポイントが低い。そんな感じなんですか？**

そうですね。それぞれの要素の組み合わせで4つ足を表していたりします。部分集合の共通の部分が「哺乳類」とかだったり、動物と植物を含むものが「生物」だったりします。

**——確かに人間の脳って、猫を猫って認識しているんじゃないって、無意識の領域で「動くもの」で「小さく」で「4つ足」で「耳が立っていて」「目が特徴的」などといった幾つかの要素を一瞬で認識して、「その要素の組み合わせは猫である」って考えているのかもしれないね。少なくとも猫という名前を覚える前の赤ちゃんは、猫をそんな感じで認識しているのかもしれないね。**

その分散表現を使うとコンピューターで概念を表現できるんじゃないかって思われてきたのですが、それがLSTMのようなリカレント・ニューラル・ネットワークでできるようになったんです。

**——なるほど、「4つ足」という表現は、猫を認識する場合でも犬を認識する場合でも使えるので、記憶しておいて、再利用できるわけですね。また再利用することで、猫と犬の概念の近さもわかる。これまでのフィードフォワード型と呼ばれるニューラルネットワークでは、毎回「4つ足」という要素を認識しないといけないけど、リカレント・ニューラル・ネットワークは、その「4つ足」の要素を覚えていて、犬を認識するときにも使うと同時に、同じ要素を持っていることで、犬と猫の概念の近さも表現できるわけです。いろいろなモノを同時に認識できるようになり、そのモノの関係性も理解できるようになった。これは確かにすごいイノベーションですね。**

例えば、「John admire Mary」という文章をコンピューターでどう表現すればいいか。これまではお手上げだったんです。自然言語処理の研究者も、「これは単語の集合で表そう、順番は無視して」という感じだったんです。

それがLSTMというリカレント・ニューラル・ネットワークを使えば、文章を分散表現で表すことができ、それをグラフ上にマッピングすると、

似たような意味の文章が近くにマッピングされる、というようなことができるようになったんです。例えば「John admire Mary」という文章なら、そのすぐ近くに「John is in love with Mary」という表現がマッピングされる。同じ単語を使っている「Mary admire John」は少し離れた距離にマッピングされます。

——admireという言葉には、「好き」「すてきたと思う」とかの要素が含まれているので、同じような要素を含む「in love with」の近くに位置するというわけですね。

そうです。今までの手法なら、こんな単純なことも扱えなかったんです。

「She gave me a card in the garden」と「In the garden, she gave me a card」が近い表現であるということも、コンピューターが理解できるようになっています。

同じようなことが、テキストだけではなく、画像でも音声でも、できちゃうようになったんです。例えば2週間ほど前に（取材日は2014年11月27日）Googleが、写真のキャプションを自動生成できる人工知能の論文を発表しました。赤ちゃんがクマのぬいぐるみを抱いている写真をコンピューターに見せれば「クマのぬいぐるみを抱くベイビー」と、非常に正確なテキストをつけることができるようになったんです。

画像を理解して、その理解に基づいてテキストを生成するという両方ができるようになったということです。リカレント・ニューラル・ネットワークの仕組みを使ってできるようになったんです。

——2年前にGoogleのコンピューターが猫の写真を自分で認識できるようになってことが驚きだったのに、もうほとんどどんなものでも認識できるようになってきてるんですね。そしてモノの関係性も理解できて、モノや関係性を文章で表現できるようになっている。たった2年ですごい進化ですね。

そうですね。例えば日本語で「私は東京に行きます」という言葉があれば、これを日本語のLSTMを使うことで分散表現に変換できるわけです。それを英語用のリカレント・ニューラル・ネットワークにかけてテキストを生成すれば「I will go to Tokyo」になりますし、アラビア語

のリアルタイム・ニューラルネットにかければアラビア語に変換されるわけです。

——これまでも機械翻訳はあったけれど、コンピューターは必ずしも意味を理解して翻訳していたわけではないですもんね。典型的な例文はきちんと訳せるけど、ちょっとでも例外的な使われ方をすると、もうお手上げ。人間は、意味を推測できるけど、コンピューターは例外には対応できませんでした。しかも言語って、例外的な表現だらけですからね。でもコンピューターが意味を理解するようになると、人間の脳同様に例外にも対応できるようになるわけですね。機械翻訳っていつぐらいになれば、使えるレベルになるんでしょうね。今はGoogle翻訳も全然使い物にならないですもんね。

いや、でもコンピューターによる英語の認識はもうかなりのレベルに達しています。ほとんどの日本人よりもうまく認識できています。

翻訳は難しいんですよ。両方の言語を知っているだけじゃなく、常識も必要になるからです。違う言語だと概念がないものがあります。例えば日本語の「しとしと雨」。英語にそれに合う概念がないんです。

——確かに「しとしと雨」って概念はないですね。僕なら、あえて訳そうとすると「静かな雨」とか「気持ちまでじめっとしそうな雨」というような表現を選びますね。

また英語の「you」には、単数の「あなた」と複数の「あなたたち」の両方の意味があります。日本語に訳す場合、前後の行間を読まないで、どちらの意味か分からない。その行間を読むためには、かなりの常識を使うんです。

——確かに。教師がクラスの生徒に「you」と言えば「あなたたち」の意味になりますが、そう訳すにはクラスには生徒が複数いるケースが圧倒的に多いという常識を持っていないといけません。そういう意味で常識が必要というわけですね。

そうです。その常識の部分も、まだコンピューターは学べていない。そこに限界があります。

あと日本語以外の言語で、翻訳精度が結構高いものがあります。コンピューターが翻訳を自動学習する上で、最適な教材が国連の記事録です。国連の記事録は同じ内容を、国連の公用語である、英語、アラビア語、中国語などの言語で書かれています。この教材を使っているのです。これらの公用語間の機械翻訳はかなり正確になっています。

あとはどれだけのコストをかけて、教材となるような大量のデータをコンピューターに読ませるかだけだと思います。英語圏の人たちは、中東や中国でどんなことを起こっているのかを知りたいのでアラビア語や中国語の翻訳にかなりのコストをかけています。試しにアルジャジーラのアラビア語のサイトの機械翻訳を、日本語と英語で比べてみてください。日本語ではなんのことか意味が分かりませんが、英語ではかなりの精度で翻訳できています。

英語圏の人は、今は日本にそれほど興味をもっていないので、日英の機械翻訳はあまり進化していません。

**——じゃあコストをかけるかどうかの問題で、翻訳のための人工知能の仕組みはだいたいできているんですか？**

知能って何だろうって話になってくると、科学者は既に知能をいろいろ定義しているんですが、ただ大部分の知能の仕組みはコンピューターで既に実現できているんです。あとは程度の問題だと思います。

**——でも人間の脳って比較的少ないデータでも、モノを認識、理解できるじゃないですか。コンピューターはまだ大量のデータが必要ですよ。人工知能はそういう意味ではまだ人間の脳に追いついていないのでは？**

現状では確かに、人間の脳に比べてコンピューターのほうが大量のデータを必要としています。ただ機械のほうが明らかに得意なことが幾つかあります。1つは疲れない、ということ。入力ソースをいくら増やしてもコンピューターは「疲れた！」と文句は言いません（笑）。もう1つは、一度学習したものを簡単にコピーできるということです。生物

は、個体の脳の状態を別の個体にコピーできない。

——なるほど。アインシュタインは、彼の脳を墓場まで持って行きましたものね。彼の脳を僕がコピーさせてもらえてたら、僕も天才になれたのに（笑）。

なので1台のコンピューターにもものすごい時間とコストをかけて大量のデータを読み込ませて賢くして、そのコンピューターが学習したことを世界中のほかのコンピューターにコピーできる。そう考えると、学習のコストってそれほど問題にならないんです。

——そうか、人間の脳、人工知能、それぞれに特性があって、どちらが優れているかということじゃなくて、それぞれの特性をうまく活用すればいい。特性をうまく利用したアプローチを採用すれば、人工知能も人間の知能にすぐに追いつくわけですね。それでもまあ今は、コンピューターは学習のために大量のデータが必要なわけですよ？

そうですね。でもそこでも大きなブレークスルーが幾つか起こっています。例えばオートエンコーダーと呼ばれる手法もその1つ。

人間や動物って、何も教えてもらわなくても勝手に映像認識できるじゃないですか。自分がどれくらい走れば、どれくらい進むのか分かっている。

——子供のキャッチボールもそうですよね。ボールがどれくらいのスピードで飛んできれば、どの辺りに着地するのかって、結構複雑な計算処理だけど、キャッチボールを練習すれば、一瞬で計算できるようになる。ものすごい映像認識力ですよ。

でも生まれたばかりの赤ちゃんはそれができないし、箱のなかで育った動物は映像認識できないという実験結果もあるんです。その学習はどのような仕組みなのか。研究者はみんな気になっていたんです。そこを解明できれば、それをコンピューターでも真似できるはずですから。



その仕組み作りを世界中の研究者が一生懸命考えているんですが、その方法の1つがオートエンコーダーです。一度画像を抽象的なレベルまで上げて、そこからこの画像を再生成するんです。再生成したものがまったく違っていれば認識が間違っている。なので1枚の画像だけ見て学習ができます。

——別の例で言うと、例えば数字の「1」の画像をコンピューターに与えれば、コンピューターはその画像の下半分を隠して、自分自身に問題を出す。脳の回路を真似た仕組みで、いろいろな回路が隠れた部分にどのような絵があるのかを予測する。下半分もまっすぐな線を予測した回路が正解なので、その回路に電気信号を通りやすくしてやる。あとは隠す部分を変えて、何度も回路に問題を出す。当たった回路の電気信号が通りやすくなるので、この自分自身への問いを何度も繰り返していると、コンピューターはこの画像がまっすぐな線であることを認識する。認識した時点でこのコンピューターに「この画像は1だよ」と教えてあげる。そうすれば、コンピューターは「1」という数字は、まっすぐな直線であると学習する。このような自分で自分に問題を出して答えていく仕組みがオートエンコーダーですよね。

そうです。動物って、こんな感じで自分自身に問題を出して答えを予測することで学習しているんじゃないかっていう考え方です。また未来を予測して、その誤差を評価することで学習しているんじゃないかとも考えられています。

——われわれは、イスに座るときに何も考えないで座るじゃないですか。イスって座るものだとして認識している。でもよく考えてみれば、イスの足が折れて後ろに転ぶ危険性もあるわけですよね。そこで座る前に「イスの足が折れるかもしれない」と予測して、恐る恐る座る。でも足は折れなかった。予測が間違っていた。そういう未来予測を何度も繰り返すことで、「イスは座ってもだいじょうぶだ」という結論に達するわけですよね。

そうです。そのほかにも、動物は自分自身が動けるので、左右に自分がかすかに動くことでモノの見え方がどう変わるのか、ということを繰り返して、そのフィードバックを受けて認識を自動調整したりもしてると

思うんです。同じ原理で、コンピューターも学習する仕組みができるんじゃないかという研究も行われています。

——なるほど。オートエンコーダーがいろいろと進化してきているわけですね。しかし2012年にヒントン教授がDeep Learningのブレークスルーを世界に知らしめたおかげで、オートエンコーダーや、ニューラル・チューリング・マシン、リカレント・ニューラル・ネットワークといったイノベーションが次々と起こっているわけですね。ものすごいスピードで人工知能が進化し始めているわけですね。そのうちに大量のデータがなくても、概念を理解できるようになるかもしれないですね。分散表現で、概念を幾つもの要素に分けて、リカレント・ニューラル・ネットワークで以前使った要素を再利用することで、1から学習する必要がなくなるのではないのでしょうか？

そうですね。ゼロ・ショット・ラーニングという仕組みの研究も盛んに行われています。ゼロ・ショットですから一度も写真を見なくても、学習できるとい仕組みの研究なんです。

——そんなことできるんですか？

オカピーを一度も見たことがなくても、「オカピーはシマウマの足で、鹿の身体している」と教えてあげれば、コンピューターが初めてオカピーをみたときでも「これはオカピーです」と答えることができる。そんな仕組みです。そんな研究が流行っています。

——なるほど。リカレント・ニューラル・ネットワークで、以前作成した分散表現を組み合わせることで、新しい概念でも理解できてしまう。人間の脳と同じですね。すごい進歩ですね。

実はDeep Learningをそれほど高く評価しない研究者もいた。匿名で話を聞いたところ次のように語ってくれた。

「(Deep Learningを開発したトロント大学の) ヒントン教授の結果は、それほどす

ごくない。彼が可能にしたのは、Back Propagation というアルゴリズム を分散計算でできるようにしたという点のみ。それ以外は何もやっていない。Deep Learning の論文をちゃんと読むと、7 段のニューラルネットワークに 1 個 1 個にフィルターが入っているが、なんでそれが上手く行くかはわかっていない。ネットワークのモデルはまだ人間が書いている。なんでもつっこめば動くというわけじゃない。多層のニューラルネットワークのデザインはそう簡単にはできない。ヒントン教授が持ち上げられているのは、業界に他にトピックがないから。従来全く解けなかった問題が解けるようになったっていうことではない」

まあ僕には専門的なことはよく分からないので、どちらの主張が正しいのかは分からない。ほかにも何人かの研究者に話を聞いてみた。

ドワンゴ人工知能研究所の山川宏所長とのインタビューを次に紹介しよう。

材日：2014年12月2日

所：株式会社ドワンゴ（東京・東銀座）

材対象：山川宏・ドワンゴ人工知能研究所所長

## ドワンゴ人工知能研究所 山川宏所長

**——カナダのトロント大学が開発したDeep Learningと呼ばれる人工知能の処理の方法について、「実はそれほどすごいことじゃない」って意見もあります。**

僕はDeep Learningという手法が出てきたことには、十分に意味があるかなと思ってます。もちろんまだ完成しているわけではないのは明らかだけど。

人間がモノを見た場合、視覚から情報が入ってきてそれを脳内で階層的に処理していて、5層か6層のところで抽象的な表現が出てくるって言われています。人間の脳の中の処理のここの部分が、これまではコンピューターではまったくできなかった。計算的神経科学の分野の人たちがずっとやってきたわけなんだけど、ずっとできなかったんです。それがDeep Learningでようやくできるようになった。初めてできたという意味で、大きな成果だと思います。長年、超えられない壁だったわけですから。そういう意味で、すごくインパクトが大きい。

——画像認識のコンペで競争項目が3つあって、そのうちの1つでしかDeep Learningはトップを取っていないという指摘もありますが。

まあようやく壁を超えたというところなので（笑）。これからに期待ですね。でもこれまで橋がなかったところに、橋がかかったということが重要なんだと思います。そこから広がる可能性が、かなりあるわけなんですから。

——人間の脳のことが全然分かっていない状態なのに、それをコンピューターで再現できるわけがない、という意見に対してはどう思われますか？

脳は電気信号で情報が伝達されているわけですから、いずれどこかの時点で脳の働きをコンピューターで再現できるようになることは間違いのないわけです。なので、今は、どの程度分かっているのかという話になるのかと思います。

これって「東京の街を知っているのか？」と聞かれるのによく似たことだと思います。長年にわたって東京に住んでいたり勤務している人は、当然「知ってるよ」ってことになる。でも「神田の街の路地裏まで知っているか」と聞かれれば「知らない」。「神田の街の路地裏のことまで分かってないんだったら、東京を知ってるっていうな」という指摘も成立するのかもしれないですが、でもまあ隅々まで知らなくても、都内の幾つかの街のことを分かっていて全体の関係が分かっていたら「東京を知っている」ということを言ってもいいのかなって思っています。

つまり何を目的としているかによって、議論が変わってくると思うんです。

神経科学者は、脳の細かな部分を調べています。薬のことを考えるには、脳の細かな部分を見ていく必要があるんです。でもニューロンって、1個からせいぜい100個くらいまでしか同時に測定できません。なのでやることは、まだまだいっぱいある。

——神経科学って、神田の路地裏を調べていかないといけない仕事。知らない路地裏はまだまだあるってことですね。そういう意味では「分かっていない」という指摘も正しい。

そうです。それに対してわれわれ全脳アーキテクチャーの研究者は、大脳新皮質が神経細胞の集団ぐらいでどんな計算をしているのかが突き止められれば、それでよしとする立場なんです。中でどうなってるかとかは、とりあえず置いておきましょうというような仕事なんです。

**——銀座は大人の街、渋谷は若者の街、新橋はサラリーマンの街。それぐらいが分かれば、神田の路地裏のことを知らなくても、東京がどんな風な都市かは分かる。地下鉄に乗れば、好きな街に行ける。そういう意味での「分かる」を目指している学問ということですね。**

そうです。そういう研究の立場からすると、特徴が分かってる部分が増え続けている状態です。

もちろん分かっている部分と分かっていない部分の違いは結構あります。割りとよく分かっていると思われるのが、大脳基底核と小脳。大脳基底核というところは、機械学習における報酬、つまり強化学習のアルゴリズムに対応する場所なんです。ここ20年ぐらいの研究で、この対応性がよく取れてきました。計算の基本原理が分かっている部類の脳の部位になります。

小脳は、メインのところは、誤差を少なくするように学習してスムーズな行動をするための装置であることが、ある程度分かってきています。

細かいところ、神経との対応づけとかは「分かっているかどうか」の議論があるけれど、おおざっぱに言って、外から見たときにどの部分で何が起きているかというところは特定されてきています。特定された部分が増え続けている状態なんです。

**——東京を知っている、脳のことが分かってきている、って言ってもいい状態になってきているということですね。では話は変わりますが、人工知能の産業的インパクトって、どう広がっていくんでしょうか？**

ちょっと前にコンピューターがプロの将棋の棋士に勝った話がありましたよね。なぜコンピューターが勝てるようになったのかというと、ある段階まではコンピューターにどう打てばいいのかということを人間が教えていたんです。ところが機械学習の技術が発達するにつれ、コンピューター

が棋譜をいっぱい読み込むことで、そこから自分で特徴量を出して、自分で打つ手を考えて打てるようになったんです。

——将棋に強くなる方法を人間が教えなくても、過去の勝負のデータをいっぱい読み込むことで、コンピューターが自分で強くなったってことですね？

そうです。コンピューターのパワーが増え、データ量も増えることで、アルゴリズムの適応範囲が広がっています。将棋の世界で起こったことが、より広い範囲で起こるわけです。

つい最近まで人間がプログラムを書いていたのに、コンピューターの処理能力が上がり、データがたまってきた瞬間に、人間を必要としない領域がだんだん増えていくんです。

——Deep Learningの成功のおかげで、人工知能の領域に研究者が集まってきたって話をよく聞きます。そうして最近、さらにいろいろなイノベーションが起こっているそうですね。その中でニューラル・チューリング・マシンの名前をよく聞くのですが、ニューラル・チューリング・マシンって何がすごいんですか？

ニューラル・チューリング・マシンはすごいです。プログラムを作れるようになったんです。まだソート（並べ替え）アルゴリズムくらいですが。でも、それができるようになった意味はすごく重要なんです。

「シード（種）AI」という言葉があります。種になる人工知能という意味です。人工知能が自分で自分を修正できるようになれば、急速に進化するはずなんです。

——増殖もできるようになりますね。

そうです。人工知能を研究する人工知能を作れば、そこから先は人工知能の進化が加速度的に速くなると思われます。今回Googleが発表したニューラル・チューリング・マシンの実装で、それがうまく行くかどうかは分かりません。多分、問題点がもちろんあると思いま

す。でも、そういう方向に向かって一歩踏み出したということは、すごいことだと思います。

自分でプログラムを書ける人工知能を作るとというのが、人工知能研究者の長年の夢だったんです。1970年代から80年代にかけては、そういう研究がいっぱいあったんです。ところがほとんどが挫折した。今回発表された論文によると、Googleのニューラル・チューリング・マシンでソートプログラムを書けるようになった。今までできなかったことができるようになったということで、意義は大きいと思います。

### ——Deep Learningの快挙で、急速に進み出した人工知能の研究ですが、このまま一気に加速して進化するのでしょうか？

まだどこかに障壁があるのかもしれないけど、今はDeep Learningが50年来の障壁を突破したところ。いろいろやり始めたところなんで、まだ次の障壁になりそうなところが見えてきていない段階です。どこかで行き詰まるのか、行き詰まらないまま進化してしまうのか。現状では分からないですね。あと2年ぐらいすると、こんなところに別の障壁があった、ということになるかもしれませんが（笑）。

### ——人工知能が進化していけば、仕事ってどうなるんでしょうか？

仕事に何を求めるか、ということに関係してくると思いますね。アメリカの組織心理学者エドガー・シャインによると、人間って8つの能力や状態のどれかが欲しくて仕事やキャリアを選択するんだそうです。その8つとは、「管理能力」「技術的・機能的な能力」「安全性」「創造性」「自律と独立」「奉仕・社会献身」「純粋な挑戦」「ワーク・ライフバランス」です。その中で、例えば「奉仕」の人は、人工知能が出てきても特に困らない。人工知能も「奉仕」、自分も「奉仕」が目的であっても競合しないですからね。一緒にやっっていこうということになります。

### ——人工知能が広く社会で利用されるようになってくると、「安全

**性」や「機能的能力」を重視する人にとっては脅威となるわけですね。**

そうですね。どういう価値観で生きていくのか、人生にとっての職業の意味をどう考えるか、というのが変化していくのかもしれないね。

**——金融市場などで、ほぼ完璧な人工知能が登場すれば、金融市場は成立しなくなりますよね？**

人工知能を独占した企業や人のところに富が集中しますね。そうなってくると、技術の問題ではなく、社会制度の問題になってくると思います。人工知能に仕事を任せる社会に向けて、どのようになめらかに移行していくのが大事です。もし本当に2045年に人工知能に人間が追い越されるのであれば、最後の10年くらいはすごい勢いで変化が起こるでしょうから、それに向けてどう備えていけばいいんでしょうか。

またそのころにはモノが足りないのではなく、精神的な充足感が足りない、ということが問題になっているかも。

**——政治はどうなるのでしょうか？**

人工知能を統括する人が必要になると思います。もちろん人工知能が世界を支配するというシナリオもありますが、多分、多くの人はその望んでいない。局所的には人工知能が正しそうな解を見つけても、全体像は人間が見てコントロールする状況を作りたい。多くの方はそう思うと思います。

ですので、人工知能の上に立つ人が必要。そういう人を政治家と呼ぶのかどうかは分かりませんが、そういう人を選挙で選ぶというのが、1つの未来像だと思います。

**——人間はICチップを埋め込んで人工知能に対抗するようにはならないでしょうか？**



チップを埋め込んでも人工知能に処理速度で勝てないと思います。律速段階が残るから。シナプス間の伝達速度以上に生身の人間の脳は速くならないんです。

## ——となると、人間は人工知能には勝てない？

人工知能に対抗するのであれば、1つの答えはマインドマップローディングでしょうね。映画「トランセンデンス」に出てきたような話。人間の脳を計算機上にコピーさせるという方法。

でもそうすると、生身の身体の方にある意識と、計算機上の意識との間に一貫性が持てないのではないか、という問題があります。そういうことを東京大学の西川先生という哲学の先生が問題提起されています。生身の身体の意識と、計算機上の意識が別のものになってしまう。それを統合しようとする、計算機上の意識が統合を拒否するかもしれないですね（笑）。

ただトランセンデンスの中にあつたような脳のアップロードは、この先に大きな技術革新がないと多分不可能でしょうね。一番必要な技術革新は、脳の中で何が行われているのかを詳細に測定できる技術が確立すること。

多分その技術が確立する前に、シンギュラリティ（人工知能が人間の脳を超える）のほうに起きて、その後すごいスキャナーを作り出すことによって、人間の脳をシリコン上に移すことができるようになるんだと思います。

## ——宇宙や心の謎の解明って、人工知能にやってもらったほうが早いという話もありますが。

人間には認知能力の制限がありますからね。人工知能だと、人間が持つ制限をかなり外せる。人間では到達できないところに到達する可能性は、確かに高まると思います。

ただ人工知能が謎を解明してくれても、人間にはそれが真実なのかどうか感覚的には理解できないでしょうけどね（笑）。

もう一人、国立情報学研究所・市瀬龍太郎准教授にも話を聞いてみた。

材日：2014年12月5日

所：国立情報学研究所（東京・神保町）

材対象：国立情報学研究所・市瀬龍太郎准教授

## 脳を模したソフトと脳を模したハードの合体で、人工知能はさらに進化する

**——人工知能が自分で自分を改良できるニューラル・チューリング・マシンという技術をGoogleが開発したって聞いたんですが。**

10月に発表された論文ですね。今回の論文のポイントは、これまでのニューラルネットワークは、そのもの自体に記憶があったんだけど、ニューラル・チューリング・マシンでは記憶が外部に作られて、内部の記憶と統合するものが作られた、ということです。

**——外付けハードディスクがついた、というような感じでしょうか。**

ハハハ。大雑把に言えば、そんな感覚でしょうかね。まあ記憶の領域ができて、それをもとに書き換えなどの作業ができるようになったんです。人間の脳は、作業記憶があって、それを統合して頭の中で推論したりするんです。そうしたことが人工知能でも、できるようになったということです。

**——じゃあ、その記憶を別のニューラルネットワークにつなげること**

## が可能になるんですか？

まあそうですね。でもまあ大事なものは、メモリのほうに書き込んだものを、あとでメモリから読みだして使えるようになった。データの読み書きができるようになったということなんです。

## ——自分で自分のプログラムを改良できるようになったわけではないんですか？

プログラムって何を指すのかにもよりますが、人工知能自身を制御するための機構自体を書き換えるわけではありません。そういう研究もいろいろなところで盛んに行われていますが、まだ完成していませんね。

## ——シードAI（種になる人工知能）の研究ですね。

そうです。シードAIは、自分で自分の制御プログラムを書き換えることで成長していくタイプの人工知能で、これが完成すると人工知能は一気に進化するのではないかって、考えられています。

今回の論文は、シードAIの観点から見ると、まだ距離があります。一步前進したことは間違いないですが、ものすごい進歩ではないですね。

## ——市瀬さんは、Deep Learningをどう評価されますか？やはりすごい進歩ですか？

そうですね。今までニューラルネットワークと人間の推論の間には乖離があったんです。人間は、思考するとき言葉を使います。しかし言葉を使わない思考の部分もあると考えられている。その差分、ギャップはどう埋まるのか。それが謎だったんです。

つまり耳や目などのセンサーから入ってくる情報と、言語による思考の間で、何が行われているのか。それが謎だった。

Deep Learningは、その分からなかった部分、ギャップを埋めてくれる非常に重要な技術です。

——思考って言語で行われていて、人間はそれ自体に気づいている。つまり顕在意識上で作業している。でも目や耳から入った情報って、潜在意識で何か処理されているはずですよ。でも潜在意識なんで、どういう処理がされているのか、自分自身では分からない。分からないので、コンピューターで再現するのが難しかったということでしょうか？

そうですね。潜在意識では、ものをパターンか何かで認識しているんだと思います。それが顕在意識になると言語という記号で認識する。パターンのレベルから記号のレベルへどうつながっているのか。それをつなげるのがDeep Learningのアプローチです。Deep Learningでそこを繋げられるようになったわけです。

画素を1つずつ認識していき、それを記号の原始的なものの単位までつなげてくれる仕組みです。

——直線や曲線という層の認識から始まって、それが集まることで目や耳という部品を認識する層になり、目や耳を含む顔が認識され、身体全体が認識される。そんな風に何層にも分けて認識を重ねていく。そんな技術ですね。それでモノの概念をコンピューターは認識できるようになったわけでしょうか？猫と犬の違いが分かるようになるには、猫がどのようなものかという概念を理解しておく必要がありますよね。その概念を理解する能力をコンピューターは得たのですか？つまり人工知能の研究者が言うところの「表現の獲得」が可能になったわけでしょうか？

はい、「表現を獲得」できるようになったのが、Deep Learningの大きな成果です。人工知能研究者の長年の課題だったので、1つの大きなブレークスルーになりました。

——そのブレークスルーを受けて人工知能の研究が再び熱くなっているみたいなんですけど、その流れで市瀬さんから見られておもしろ

## ろい動きに、どんなものがあるのでしょうか？

ハードウェアもおもしろいものが出始めたということが1つありますね。脳の仕組みを真似たICチップをIBMが開発しています。シナプスチップと呼ばれるものです。

今までは汎用のコンピューター上で人工知能を作ろうとしていたのですが、これからはIBMのシナプスチップのように脳を模したハードウェアで、しかも人間に近いような処理速度のものが出てくるわけですから、これから人工知能がますますおもしろくなりそうです。それがどういう形で収束していくのかは、まだちょっと分からないですが。

**——でもIBMのシナプスのチップには、Deep Learningの仕組みは載ってないですね。**

はい、今のところは機械学習の機能は入っていないですね。既に学習済みのものを入れているみたいです。でもいずれ機械学習の仕組みもシナプスのチップに搭載されるようになるのだと思います。

**——まあ、そうでしょうね。脳を真似たハードウェア以外では、ほかにもどのような動きに注目されていますか？**

あとは特定の脳の機能ではなく、脳全体をコンピューターで作ってしまおうという動きが興味深いですね。

ヨーロッパとかアメリカでは、脳の中身全部をシミュレーションしようという動きがあります。

**——人間の脳を全部スキャンして、それを再現しようという動きですよ。fMRIとかの性能が急速に向上しているので、脳を全部スキャンするのが夢物語じゃなくなってきたって聞きます。またナノロボットを脳のシナプスに貼り付けて、どんな動きをしているのか逐一調べるって方法もあるみたいですね。**

ええ、ただ脳の中身自体を全部シミュレーションするのって、やっぱり

大変なんです。脳には1000億個のニューロンがあると言われてます。それらの動きを全部把握するには、計算機のパワーが必要です。またシナプスの中には、なくても脳全体の機能にほとんど影響を与えないものも多いんです。

なのでわれわれ日本人研究者は、もっと大雑把というか、大枠で物事をとらえてもいいんじゃないかって思って研究を続けています。つまり脳のそれぞれの部署の機能を理解し再現し、それを組み合わせることで脳と同等のものを作れるのではないか、とっているわけです。全脳アーキテクチャと呼ばれるプロジェクトですが、日本はそういうアプローチでやっています。

——例えばある研究者は海馬を担当し、別の研究者は大脳新皮質を担当する。それぞれが作った仕組みを合わせることで、脳全体と同じようなものを作ろうというプロジェクトですね。なるほど、欧米とは別のやり方で研究を進めているわけですね。

## 人工知能の進化は人間の脳の進化？ 脳にチップを埋める必要なし

——脳を模したソフトである人工知能が、脳を模したハードであるIBMのICチップ「シナプス」に搭載され、それがいろいろな機械に搭載されていく。当然ロボットにも搭載されていきますよね？

そうですね。最初は監視カメラなどの画像認識系のところから入っていくのですが、最終的にはロボットに乗って行って現実の世界とオーバーラップしたような形になっていくのかもしれないですね。もしくはロボット自体も必要でなくなる世界になるかもしれないですね。

——え？ロボットが必要でなくなる？どういうことなんでしょうか？

人間が目で見えるものと、バーチャルなものが重なっていくと思うんです。例えばGoogle Glassなんかそうですね。それがさらに先に進めば、人間の脳自体に何らかの作用をして、直接コンピューターとやり取りする世界もありえるかもしれません。そうなればロボットと会話して、ということも必要なくなる可能性もあると思います。

——高度な人工知能を搭載したロボットに何かの計算処理を頼まなくても、人間の脳自体が高性能になっていく可能性があるということですね。人間の脳が進化すれば、人間を超える人工知能は確かに不要になりますね。ロボットは、あくまでも人間の言うことを聞くだけの機械になります。人間と対等、もしくは人間を超える知能を持ったロボットは不要になりますね。でも僕は脳の中にICチップを埋め込むというやり方には抵抗があるんですけど。

確かに抵抗がある人も多いかと思います。今までは脳の電気信号を読み取ったり、信号を与えたりするには、チップを埋め込むしかなかったんですが、でも今は脳の解析技術がすごい勢いで進歩しています。脳をスキャンすることのできるfMRIなどの機械は、今はものすごく高価で大きなものが使われていますが、それも低価格化、小型化が急速に進んでいます。

——帽子の中にfMRIが搭載される、なんてこともありえますか？

すぐに実現することは難しいですが、可能性はあります。その辺りのことは、現在の技術の延長線上で考えるというより、技術が加速度的に進化するという前提で考えてみると、この先、そうした技術は、ものすごいものになる可能性は十分にあります。

——そうすると、人間の脳にチップを埋め込まなくても、人間の脳とコンピューターがつながる可能性もあるわけですね。

はい。今でも脳を手術で開かなくても、fMRIなどで脳の中で何が起

こっているのが、ある程度は分かるようになってきています。そうした技術がすごい勢いで進展しているわけですから、いずれ人間の脳に電極を埋め込まなくても、人間の考えていることが分かるようになる可能性はあります。

例えば、サイバーダイン株式会社が開発した、人間の動きを補助するロボットスーツなんかは、筋肉の表面から電気信号を取得しています。脳に電極を埋め込む必要がないわけです。

### ——皮膚の上から神経を通っている電気信号を読み取れるもの なんですか？

読み取れます。ただとても小さな信号なので、解析技術がしっかりとしていないとだめですが。それでもかなり取れるようになってきています。また電気信号だけではなく、皮膚の動きなどの信号も取って、人間の動こうという意思を、総合的に判断しているみたいです。

### ——脳波でも脳の動きが読み取れますよね。

脳波でもある程度のところは分かるんですが、ただ雑音が入るんです。脳の特定の部分の動きをピンポイントに取るという技術の研究、開発は、最近あちらこちらで行われていますが。ただやはり今のところは、ざっくりした情報になってしまい、たいしたことが分からないんです。

### ——人間の脳の信号を読み取れたり、反対に脳に信号を書き込んだりということが、電極を打たずにできるようになれば、人間は生身の身体のまま、脳を進化させることができるわけですね。そうになると、ロボットだけが進化して、人間より優秀になるということもなくなりますね。話は変わりますが、人工知能の研究を国策プロジェクトにすべきだと思いませんか？

そうですね。大きな装置が必要だったりしますからね。Googleなんかは、日本の大学が研究にかけている予算とは桁違いのお金をかけて研究を進めていますからね。そこと伍していこうというのであれば、現



状では限界がありますね。

**——中国も人工知能に力を入れていますしね？**

そうですね。どこが一番先にとるか、みなさん争っているところですね。

**——人工知能って社会にとんでもない影響を与える可能性がありますよね。一步間違えば、原子力と同じように悲惨な結果を生みかねない。社会として人工知能とどう向き合っていくべきかというような議論って、既に行われているのでしょうか？**

私は人工知能の技術は、人間と対峙する技術ではないと思っています。あくまでもツールでしかなく、人間が幸せに過ごすためのものだと思います。ただ使い方を間違えると不幸な結果になると思います。

人工知能をだれが持つのかという議論も必要になってくるでしょうね。軍が持つのがいいのか、民間企業が持つのがいいのか。いろいろな議論が出てきています。

また人工知能が財を作るようになって、人間はその財の恩恵を受けただけでいい世の中になったとき、人間の存在意義はどうなるのか、という問題もあります。

人工知能って、技術論だけではなく、人間を含めた社会すべてに影響を与えるものとして、考えていかないといけないと思っています。

今、ちょうど哲学や倫理の研究者の方たちと研究会を作って、そうしたことを議論し始めたところです。

**——人工知能がさらに進化していけば、宇宙の謎も解けるようになるのではないのでしょうか？**

そういうふうに考えている人もいますね。今までの進化の歴史を考えれば、猿から人間に進化したけれど、この先は人間よりかは機械が進化の頂点を極めて、それで宇宙に出て行くのではないか、というように

考えている方もいます。

## 動き出した日本のプレイヤーたち

研究者とのインタビューを並べたが、このように人工知能が急速に進化するのであれば、そこにはビジネスチャンスもあるはず。ビジネス領域でのキーパーソンにも取材した。その結果を、ブログ記事「人工知能時代の覇権はオレたちが獲る 動き出した日本のプレイヤーたち」として公開した。以下にそれを掲載しよう。

### ◎人工知能時代の覇権はオレたちが獲る 動き出した日本のプレイヤーたち

脳を模した人工知能の研究領域に研究者が集中し始めたことで、人工知能が急速に進化し始めている（関連記事「人工知能が急に進化し始めた」）。テクノロジーが急速に進化すると、テクノロジー業界の勢力図が塗り変わる可能性がある。スマホ＋クラウドの次のパラダイムは、だれが覇者になるのだろうか。

<http://thewave.jp/archives/1985>

### ▶Google、Facebook、中国Baiduも

次のパラダイムの覇権に最も近い位置にいると見られているのがGoogleだ。同社は、2014年1月にイギリスの人工知能系ベンチャーDeepMind社を推定5億ドルで買収したほか、10月にはオックスフォード大学発ベンチャーのDark Blue Labs社とVision Factory社の2社を買収し、両社の主要研究者7人をDeepMind社に移

籍させている。

DeepMind社はどれだけすごいのか。同社の技術に触れたカリフォルニア大学スチュアート・ラッセル教授をして「DeepMind社の技術は、私たちが考えていた技術的レベルをはるかに上回る予想外のもので、大きなショックを受けました。私を含めて多くの人が思考停止に陥ったと思います」と言わしめるほどだ。（関連記事 <http://gigazine.net/news/20141203-deepmind-demis-hassabis/>）

DeepMind社の買収をめぐるのはFacebookがGoogleと最後まで競い合っていたと言われるように、Facebookも人工知能に力を入れている。Facebookは昨年末にニューヨーク大学のYann LeCun教授を雇用し、人工知能研究所を設立している。同研究所は、Facebookの本社のあるカリフォルニアと、ロンドン、それとニューヨーク大学のすぐ隣の3カ所に拠点を設けている。

そしてIBMも「人工知能に社運を賭けている」（同社関係者）というほど力を入れているし、「Microsoftもこの分野には古くから力を入れている有力企業です」と、この領域に詳しいビジョン&ITラボの皆川卓也氏は言う。

また中国企業も人工知能に力を入れ始めた。中国のネット検索大手Baidu（百度）は約3億ドルを投じ、5月にシリコンバレーに人工知能の研究所を開設した。昨年、北京に開設した人工知能研究所と合わせて、人工知能研究者として有名なスタンフォード大学のアンドリュー・ング氏が統括するのだという。

[http://www.nikkei.com/article/DGXNASGN1700N\\_X10C14A50](http://www.nikkei.com/article/DGXNASGN1700N_X10C14A50)

日本企業の動きはどうなんだろう。某テクノロジーコンサルタントは「日本の大手企業の動きを、ウォッチする必要はまったくありません」と言い切る。

人工知能に関しては日本も「第5世代コンピューター」という国策プロジェクトがあったおかげで研究者の層は厚いのだそうだが、問題は研究予算。国立情報学研究所の市瀬龍太郎准教授は「Googleとかは日本の大学とは桁違いの予算をかけて研究をしています。そこ

と伍していくのには、どうしても限界があるように思います」と語っている。

早稲田大学政治経済学の井上智洋助教は、雑誌エコノミストに「政府は人工知能の研究開発の支援を」という記事を寄稿している。人工知能は社会のインフラ的存在になり、日本の生産性を劇的に上昇させる可能性があるから、というのがその理由だ。

### ▶データを集める仕組みが大事

こうした状況で、果たして日本は、英、米、中国に対抗できるのだろうか。

ただ基礎研究で後塵を拝していても、実際のビジネスで影響力をつかめればいいという意見がある。「大事なのは、人間に対するユーザーインターフェースの1枚目を取ることなんです。2枚目、3枚目と距離ができるほど収益は小さくなる」とソフトバンクロボティクス株式会社の林要（はやし・かなめ）氏は言う。

インフラを押さえた方が強いのか。ユーザーに近い場所を押さえた方が強いのか。よく議論になるテーマだ。

ただ携帯電話事業に関する限り、インフラとなる通信網を提供するキャリアの状況は厳しくなる一方だ。各社とも同じような携帯電話端末を扱っているので、競争できるのはサービスの質と価格だけ。どうしても価格競争で、収益を圧迫する結果になってしまう。「キャリアは単なる土管になっていく」。そんな意見が業界関係者から聞こえてくる。

一方で、ユーザーと直接接点のある企業は強い。最近ではLINEがモバイル決済サービス「LINE Pay」を始めるなど、ユーザーとの接点を活かして新たなサービスを次々と生み出している。ユーザーに一番近いところ、林氏の言う「1枚目を取っている」ことからくる強みだ。

「Appleにしろ、Google、Facebookにしろ、みんな1枚目を取る競争をしているんです。ソフトバンクも今度こそ、そこを取りに行くべきじゃないかって考えています」と林氏は言う。

検索の「1枚目」はGoogleが手にした。ソーシャルはFacebook、スマートフォンはApple、Googleが手にした。人工知能ではソフトバンクが「1枚目」を取りたい。ソフトバンクが、世界に先駆けて人型ロボットPepperを家庭向けに本格投入しようというのはこのためだ。

Pepperは人の心を癒やすコミュニケーションロボット。人を喜ばせるにはどんな情報がいいのか、どんな会話がいいのか、というデータがPepperに蓄積される。またPepper一体、一体から送られてくるそうしたデータは、クラウド上で集計される。それをベースにクラウド上の人工知能が順調に学習を進めていけば、他社に真似のできない人工知能が出来上がるはず。この人工知能を作るために、破格の価格でPepperを発売しようとしているわけだ。

米IBMはソフトバンクロボティクスと提携して、Pepperのデータ処理に力を貸すことになった。「データがないとIBMの人工知能は賢くならない。IBMはそのことを理解しているんです」と林氏は言う。

東大発の有力ベンチャー、プリファードインフラストラクチャー社の岡野原大輔副社長も、データを集める仕組みを持つ少数の企業が影響力を持つようになると思う一人だ。

「今のネットもオープンな仕組みとは言われますが、Googleが検索ユーザーの動きをほとんど把握しています。Googleを使えなかったら、ほとんどの情報にリーチできない状態です」「同様にこれからも、オープンだけど事実上そこを全部コントロールする仕組みができるのではないかと思います。そこにはネットワーク効果が働くので、強いところがどんどん強くなると思います」と岡野原氏は語る。

次のパラダイムでは、少数の覇者がすべてを牛耳るようになるというわけだ。

具体的にはどのような仕組みなのだろうか。他社との秘密保持契約があるらしく岡野原氏からは多くを語ってもらえなかったが、ネットワークにつながった無数のデバイスから集まったデータを人工知能が取り込むことで、人工知能がどんどん学習していく仕組み作りを考えているようだ。

そして、あらゆるデバイス、あらゆるアプリの後ろには人工知能がつかがるようになる。そのとき「中心になるのは、今の大手ではないと思いますよ。新たに勃興してくる会社が新しい時代の中心になるのだと思います」と岡野原氏は予測する。「過去のコンピューターの歴史を見てもパラダイムが変わるたびに新しいプレイヤーの時代になっています。今のスマホ、クラウド時代の覇者は、今の成功に縛られて新しいことができないんじゃないかと思います」。

次の時代を見据えて、勇者たちが動き出した。ワクワクする時代の幕開けだ。

#### 【蛇足、オレはこう思う】

人工知能に関して、何人かのキープレイヤーを取材してみて、おもしろいことに気づいた。次の覇者として、だれもAppleの話をしていないのだ。もちろんスマホ時代を築いたという功績に対してのレスポンスはかなりのものなのだが、人工知能の覇権争いの話をしていても、だれもAppleを意識していない。

Googleの話は出る。しかしGoogleにしても、それほど神経質にはなっていない模様。というのは、買収は繰り返しているものの、Googleとしてまだ新しい動きが出てこないからだ。最近では、Googleのロボットの責任者が退社したというニュースがあった。Google内部で何が起きているんだろう。「有力ベンチャーを買収することで天才を数多く雇用したんだけど、天才は協調性がないことが多いので、それぞれがバラバラに動いてまとまりが取れないのではないのでしょうか」。そう分析してくれる関係者もいた。

業界関係者の間では、家庭用ロボットではまずはGoogleが動くと思われていたのに、ソフトバンクに先を越された。このととも、原因がその辺りにあるのかもしれない。まったくの憶測だけど。

やはり、新しいパラダイムは、新しいプレイヤーが中心になるのだろう

か。

ソフトバンクの林氏の見解については、後で詳しく述べるとして、岡野原氏はどのように語っているのか。ビジネスに関する岡野原氏の取材メモの続きを次に掲載しよう。

**——画像、テキストの処理に人工知能が使われるようになるでしょうが、特に映像の部分がコンピューターで理解できるようになると、いろいろな仕事が機械に置き換わるのではないのでしょうか。**

特に監視カメラ、防犯の領域ですね。今の監視カメラって、警備員はほとんど見ていないじゃないですか。

**——まあずっと監視カメラのモニターとにらめっこするのって大変ですからね。**

今だと、監視カメラの前で人が倒れていても、後から気づくということのほうが多い。でも監視カメラに人工知能を搭載すると、何が起きているのかをコンピューターが把握し、異常事態を教えてくれるようになります。

**——今は事件が起こってから確認のために監視カメラを人力で見直すという感じで使われていますが、事件が起こっているときに警告を発信してくれるようになるわけですね。**

あとは機械を無数につなげて、協調して動かすこともできるようになります。カメラは今、すごく安くなっているので、あちらこちらにばらまいて、つなげて、何が起きているのかを理解する万能センサーができるんじゃないでしょうか。

——なるほど。ディズニーランドのアトラクションの待ち時間予測も、よりリアルタイムで正確になりますね。交通渋滞予測も。自動車にもカメラが搭載されて、自動車で情報を共有するようになるでしょうし。

そうですね。あとは店舗内やこれまでカメラが置かれていなかったところにも置かれて、何か異常があれば対応できるようになる。

——無人自動車を好きなときに呼び出して、乗り捨てられるようなカーシェアリングとか。

あとはコンビニの無人化も、ものすごい勢いで進むかも知れないですね。ど的人がどの商品を手にとったのかもカメラで分かるので、レジに並ぶこともなくなります。万引きもできなくなる。おかげで今までコンビニがなかったような新興国や過疎地にも、無人コンビニができるようになるかもしれないですね。

特に新興国での変化がおもしろそうです。進化は新興国で先におこるのではないのでしょうか。

——リバーズ・イノベーションと呼ばれる現象ですね。抵抗勢力もないですし、技術の進歩が早い場合は、既存の仕組みを改良するよりも、1から作ったほうが早いですからね。御社としては、どの領域を攻めていこうと考えているのですか？

われわれPreferred Networksのミッションは、IOT（モノのインターネット）の時代に、ネットワークに知能を埋め込むことです。例えばカメラのネットワークを人工知能につなげて、カメラが協調し合って何かを自動的に認識して追いかけたり、自動的にアクションを起こすような仕組みを作りたいと思っています。例えばデジタルサイネージに出す広告を変える、などといったことです。

——やはり人工知能に注力するわけですか？



そうですね。今、京都大学のバイオのグループとゲノム解析の研究で協力しているんですが、化合物の活性予測とかでDeep Learningを使っているんです。そこでもDeep Learningはすごくうまくいってます。

**——やはりこれからのアプリやIOTのようなモノはすべて人工知能につながってくるのでしょうか。**

そう思いますね。

**——いろんなものがつながった人工知能って、だれがコントロールするようになるのでしょうか？そこをコントロールする人や企業、国家に、大変なパワーが集中するように思うのですが。**

そこは僕もすごく興味があるところです。人間には、自分が持っている知能を他の人間と共有するために言語があります。コンピューターも同様に、知能を交換、共有するような仕組みができてくるのだと思います。それは単純なプロトコルとかのレベルではなくて、人間の言語並みに拡張性、柔軟性があって、相手が知らないこともちゃんと説明できて、というようなものができあがっていくのだと思います。

**——時代の進み方としては、最初はIBMやGoogleといった大手企業の人工知能を利用する時代があり、それから、もっと分散的で、オープンで、フラットな時代へと移行していく。そんなイメージでいいのでしょうか？**

そう思いますね。ただ多分、方言のようなものは残ると思います。人間の言語でも、日本人同士では日本語で話すけれど、日本人の中には英語ができる人もいるので、その人を介して外の世界ともつながっています。同じように、IBMの人工知能を利用するコンピューターはIBMの言葉を使うけれど、Googleの人工知能と話をするときには別の言語を使ったり、翻訳したりする。そんなイメージですね。

——権力が少数に集中しないほうが理想的な世界だとは思いますが、現実問題としては、少数の大企業に影響力が集中してしまいそうですね。

その可能性は結構あると思いますね。今のインターネットもオープンだと言われます。確かにサーバーは分散されていますが、ほとんどの人がGoogleを介して情報にアクセスしている。Googleは検索ユーザーの動きを全部把握していますし、われわれユーザーはGoogleを使えなかったら、ほとんどの情報にはリーチできないわけです。

同様にこれからも、オープンなプラットフォームだけど、事実上そこを全部コントロールできるようなサービスか仕組みができるのではないかと考えています。そこにはネットワーク効果が働くので、大きいところがどんどん強くなって、そこがかなり支配的になる。一部分はそれに敵対するような動きもあるかもしれないが、大半の人は「便利だからいいか」と思うようになると思いますね。

——やはり寡占化は避けられないということですね。

あともう1つ思うのは、中心になるのは今の大手ではないということです。新たに勃興してくる会社が、新しい時代の中心になるのだと思います。

——それはどうして？

今までのコンピューターの歴史を見ても、パラダイムがシフトするたびに、新しいプレーヤーの時代になるからです。スマホ、クラウドの時代が終わりIOTの時代になると、スマホ、クラウドの時代の覇者は今の成功に縛られて新しいことができなくなる。マイクロソフトなんかはそうですね。

——成功した企業って1つ前のパラダイムで十分儲かるから、新しい領域にはそれほど本気に取り組まないですものね。ということは

**スマートフォン、クラウドの今の時代の覇者であるAppleやGoogleは、次のIOTの時代では成功しない？でもGoogleって、斬新な動きをするじゃないですか。**

そうですね。Googleはちょっと特殊ですね。（前CEO）のエリック・シュミット氏から（現CEO）のラリー・ページ氏に経営トップが変わってギヤが変わりましたね。インターネット、PCの時代からモバイルの時代、さらにIOTの時代に向けて変わろうとしているところですね。それでIOTに関わるような会社を次々と買収して、ポストモバイル、ポストクラウド時代の優秀な知能を外部から取り込んでいますね。

内部でイノベーションを起こすのって、難易度が高いんですよ。自分たちのビジネスを食ってしまう可能性があるんで。なので、外部から取り込むか、別動部隊にするしかないんです。その別動部隊が（秘密研究所）「Google X」。Googleは次の時代に合わせて、非常にいい位置にいると思います。

## IoT x 人工知能で、人間の脳の限界を超える

岡野原氏のお話があまりに面白かったのでプリファード・ネットワークスの代表の西川徹氏にも話を聞くことにした。取材したのは2015年1月14日。この電子書籍の中で、最も旬な情報だ。

宇宙、人の心、生命の神秘。科学ではまだまだ分からないことが山のようにある。科学でまだ理解できないのは、人間の認知能力に限界があるからかもしれないし、対象が複雑過ぎるからかもしれない。人間には、未来永劫解くことのできない謎なのかもしれない。

だが複雑なものの理解が得意な人工知能なら、それらの謎を解けるかもしれない。東大発ベンチャーのプリファード・ネットワークス（PFN）の西川徹氏は「人間の脳って優秀で、1台の機械では超えられない部分がまだまだあるんです」とした上で、「でも、無数のデバイスを1つの神経系のようにつなげれば、人間の能力を超えることもありえるかもしれない」と語る。

モノのインターネット（IoT）に人工知能を掛け合わせることで、今までにないような価値を、いろいろな領域で実現できる。西川氏はそう指摘する。自動車に搭載することで事故のない社会を実現できるかもしれない。身体データを採取し人工知能で解析することで生命の謎の解明につながり、医療が大幅に進化するかもしれない。宇宙の謎だって解けるかもしれない。

こうした謎の解明のために、それぞれの領域から支援を求められているのがPFNだ。シリ

コンバレーの後追いではなく、世界の最先端の技術の担い手になりたい。その壮大なビジョンを西川氏に語ってもらった。

取材日：2015年1月14日

場所：株式会社Preferred Networks（東京・本郷3丁目）

取材日：同社・西川徹代表取締役

## ▶世界の最先端を狙うために新会社を設立

**——プリファード・インフラストラクチャー（PFI）と、プリファード・ネットワークス（PFN）の2社を運営されていますが、この2社の業務内容の違いを教えてください。**

分かりました。まずはこれまでの経緯をお話しますね。僕はプログラミングが好きで昔からプログラミングコンテストによく出ていたのですが、大学院のときに世界大会に出場したんです。そのときに同じくその大会に出場していた京都大学の学生たちと仲良くなった。このメンバーで何か一緒にできればおもしろいよね、という話になったんです。

そのころ、岡野原（副社長）ともう一人の創業者で、東大の近くにあったバイオベンチャーで働いていました。そこではsiRNA（低分子二本鎖RNA）の解析プログラムを岡野原らと開発したんです。そのときに「ベンチャーを作ってみるのもおもしろいかな」と思って、PFIをスタートさせたんです。

最初は開発力を高めようということで、分散検索エンジンを手がけました。岡野原が未踏プロジェクトで作っていたプログラムがあったので、それを元に、その上に分散システムを構築するカタチで作りました。そしてそれをレコメンデーションエンジンに拡張したりとかしていました。

その流れで途中から「Hadoop」のコンサルティングを行うようになりました。（編注：Hadoopは、Googleの論文をもとにオープンソースとして開発されている分散バッチ処理ソフトウェア）。

それを1年間やったときに、Hadoopは自分たちで作ったソフトではないので、自分たちでソフトを作りたいということになりました。そのときにNTTプラットフォーム研究所（現ソフトウェアイノベーションセンター）からお声がけをいただいて、何か共同でプロジェクトを進めようとい

う話になりました。当時、我々は機械学習に一層注力しておりましたので、機械学習を超分散環境にもスケールさせられるようにしたい、という提案を行い、一緒にJubatus（ユバタス）というオンライン機械学習向け分散処理フレームワークの開発をスタートしました。これはオープンソースのソフトで、機械学習をたくさんのコンピューターで分散並列で動かす仕組みです。

それまでは国内の活動が中心だったのですが、このソフトを開発したこともあって、これからは世界でも活動しようということになり、USへの進出も検討し始めました。

それで何回かUSへ出張したのですが、あるときネットワーク機器メーカーのシスコ・システムズのフェローと会うきっかけがあり、彼にJubatusの説明をしました。シスコはフォグ・コンピューティングを推進していました。（編注：クラウド・コンピューティングは中央のコンピューターでデータを処理するが、フォグはより現場の近い部分で処理をするという考え方。中央にまでデータを転送しなくていいので、データ転送が迅速になるなどの利点がある）

われわれは、データをネットワークの真ん中にためるのではなく、エッジで処理する「エッジ・ヘビー・コンピューティング」を提唱し始めていたので、シスコのフォグ・コンピューティングと似通ったコンセプトであるということもあり、シスコのフェローと意気投合しました。そこで僕らのJubatusが、彼らのフォグ・コンピューティングの頭脳になるのではないかと思って協業を検討することになりました。それで、子会社をUSに設立したわけです。

そのころ日本では、IoTという言葉を使う人は、まだほとんどいませんでした。でもその後、IoTが一気に脚光を浴びるようになってきました。でもそれでもIoTに機械学習を使っていこうという流れは、世界でもまったくありませんでした。

僕自身は今まで自然言語処理を中心にやってきたんですが、今まで作ってきた技術はどうしてもGoogleの後追いになっていました。なんだかそれが嫌だった。でもIoTに機械学習を応用していくという領域は、世界的に見てもまだだれもやっていない。この領域なら自分たちが最先端の技術の担い手になれる。世界の最先端を走りたい。そういう気持ちが大きくなってきたんです。

PFIは外部出資を受けないというポリシーで運営してきました。ですがそのポリシーを捨てても、一気に加速してわれわれの技術を普及させて行きたいと思うようになりました。そこで、1年前にPFNを作りました。私は、両方の社長をしていますが、活動の主体は今はPFNになっています。

## —— 2社の業務内容などの住み分けはどうなってますか？

PFIは出資を受けないというポリシーのまま、検索エンジン、自然言語処理を中心になっています。人が生み出したデータを解析する領域です。

一方のPFNは出資を受けるといポリシーで、マシンが生み出したデータを解析する領域です。ビジネスを最大に加速させることを目指しています。フォーカス分野はIoT。IoTを支えるためのプラットフォームを作っていくつもりです。

## ——IoT x 機械学習は、世界的に見てもだれもやっていない領域ですか？

だんだんする人が増えてきています。Googleの自動運転もいわば、IoT x 機械学習ですからね。Googleとかは、この領域に力を入れ始めてますね。

## ——でもまだ覇者はいない？

まだいないですね。

### ▶ 人間の脳のキャパを超えるIoTデータの量と多様性

## ——IoT x 機械学習は何が難しいのでしょうか？どこがキモになりそうですか？

人が生み出すデータというものは、たかだかしれてるわけです。人口 x 一人当たりのデータ生産量なわけですから。でもIoTのほうは、デバイスはいくらでも増えるし、デバイスが生産するデータも解像度をいくらでも高めることができる。なのでデータ量は、ものすごく増えることにな

ります。

航空機が取得するデータでも、1回太平洋を横断するだけで、人が一生かけて生み出すぐらいのデータを生成するわけです。あまりのデータ量なので、今はそれをリアルタイムで通信できていない状態です。もしそれを全部通信できるようにすれば、ものすごい量のデータが集まるわけです。

IoT x 機械学習の難しさの1つは、データのスループットが増えてきているという状況があるということです。

もう1つの難しさは、データが多様になってきているということがあります。

今の機械学習の対象となるデータは、映像データが中心です。人の顔を認識するといったようなタスクに機械学習が使われることが多いわけです。でも今後IoTが発展していくと、いろんなデバイスがわけの分からないデータを生み出し始めます。

例えば生体データです。生体データを採取しても、それがどういう意味を持つのか、なかなか分からない。そういうデータの周辺にさらに多くのデータが生成されてきています。

そういうデータを見ても、それが何を意味しているのかは人には分からないんです。

これまでのデータ分析は人が行ってきましたから、人が理解できるデータばかりを収集してきました。データサイエンティストたちが統計学を駆使して、データの意味を汲み取っていたわけです。

しかしIoTではいろんなデータが取れるようになってくる。多くのデータは人が見ても、意味を汲み取れないものになってきます。意味を汲み取れるだけの知識を、人間が持つことに限界があるわけです。人がデータ解析のボトルネックになってきているのです。

データ量の増加と、データの種類の増加。この2つが理由で、人では対応できなくなっています。

ちなみに、人が生み出すデータの中でも最も解析が難しいのは自然言語だと思っています。自然言語を機械が本当に理解できるようになるのは、まだまだ先の話ではないでしょうか。

——何年くらい先になりそうですか？

20年くらいかかるかもしれませんね。

今の人工知能の技術であるDeep Learningの性能は、まだまだ動物の視覚神経ぐらいの能力しかありません。人の脳にはまだ近づけてはいる。人の脳は、まだ中で何が起きているのか分からない。試行錯誤がこれからも続きます。

私感も入るけど、人間ってコミュニケーションする際には言葉以外の情報も吸収していると思うんです。例えば映像情報。日々、周りを見渡すことで、いろいろな情報を得ている。例えば、「イスは空を飛ばない」という概念を日々の生活の中で掴んでいる。言葉で説明を受けなくても、生活の中で常識としてつかんでいるんです。そうした情報、自然法則のようなもの、言葉にされていないコンセプトなども、自然と解釈した上でコミュニケーションしている。そういった言語以外の情報、概念、常識をも理解しないと、本当の意味で言語を理解できないと思います。

今は、大型コンピューターに自然言語の文書を大量に読み込ませることで、コンピューターがある程度、自然言語を理解できるようになります。本当に言語を理解できるようになるためには、コンピューターは言語以外の情報をも理解しないと無理だと思うんです。

なので最終的にはIoTと自然言語処理コンピューターが結びつき、そうなることで本当に人工知能が言語を理解できるようになるのではないかと思います。

いろんなセンサーデバイスを使って環境で何が起きているのか、より精緻に理解しようというのがIoT。そこでデータがたまり理解が深まると、世の中にある言語化されていないコンセプトをたくさん蓄積できる。そのデータベースと言語のデータベースを突き合わせることで、人工知能がより人に近づくことができると思います。

## ——世の中の言語化されていないコンセプトにはどんなものがあるのでしょうか？

例えば色のコンセプトは国によって異なりますよね。虹を7色という文化もあれば、5色という文化もある。文化が違えば中間の色には、名前がついていないんです。

潜在意識下では、恐らくいろんな色を認識しているはず。中間色に反応する細胞があるはずなんです。深い層の中にはいろいろラベルつけられていないものもあると思います。

でもラベルつけられていないものが役に立っていないのかというと、僕は役に立っているんじゃないかと思っています。



——うまく言葉で表せないけど、なんとなくこうするのがいいと思う。直感がわく。そんなことがありますよね。そしてそうした直感のほう  
が、判断が正しかったりする。

そういうことです。言語化されていない部分に、重要な情報があると思うのです。

▶工場、店舗、防犯で既に実証実験中

——なるほど。ところでIoT x 機械学習のビジネスの事例には、どのようなものがあるのでしょうか？

直近だと、日本では工場のオペレーションの最適化にIoT x 機械学習を利用しようとしています。人が工場の中でより効率的に動けるように、オペレーションの最適化を目指すわけです。

今までは工場の作業の流れを、人が考えて、それに基づいて工場内の機械の設置場所を決めていたわけです。その部分をIoT x 機械学習で最適化しようと考えています。具体的には、工場をセンサーで監視し、どういう風に人を配置したり動かしたら、生産の効率を最適化できるか。それを機械学習を使って、つかんでいこうとしています。

——センサーにはどういうものがあるのですか？

今はカメラですね。映像データの解析に力を入れています。カメラを配置して、だれがどの機械にアクセスしているのかを解析し、無駄に人が滞留しているところを見つけて改良する仕組みです。

また製造ラインの監視も映像で行って、問題が発生しても自動認識し、そこにカメラを向けてフォーカスを当て、制御ルームから問題を解決できるような仕組みもあります。

**——なるほど。製造業にIoT x 機械学習が応用できるわけですね。ほかにはどんな事例がありますか？**

USでは、小売業の店舗内でのお客さんの動きを解析しています。

これまでの小売業は、レジ情報しかなかった。客が買ったという結果しか分からないわけです。でも実際には、入店から買うまでに、いろいろな動きがありますよね。商品を手にする。「微妙だな」と首を傾げる。同じような商品を手にとって比較する。こうした客の動きの情報は、店舗側にとっては非常に価値ある情報のはずです。こうした情報を解析する仕組みを開発しています。

**——棚ごとにカメラを設置するのですか？**

そういうアプローチもあるでしょうし、高解像度のカメラを天井に設置して店内を見渡す、という方法もあるでしょうね。

**——ほかにはどんな事例がありますか？**

日本では、防犯も手がけようとしています。ビルのセキュリティカメラの監視は、今は人がやっています。カメラを見ていて、事件が起きてから警備員が駆けつけるわけです。

そうではなく、コンピューターが映像を自動解析し、問題が起こりそうだったら何らかのアクションを起こすことが可能になります。例えば、自動的にドアを閉めるとか、万引きしそうな人に対して「万引きは犯罪です」というようなメッセージを表示して抑制する。そういったアクションです。

**——この3つの事例って、もう動き出しているのですか？**

3つの事例のうち、小売りと防犯は実証実験が始まっていて、工場最適化はまもなく実証実験が始まります。

**——こうしたソリューションは自社で開発して、自社で営業して販売するのですか？**

いえ、防犯は某警備会社と組んで開発を進めています。こんな技術がほしいというお話は、大手SIer（システム開発会社）からいただいています。われわれはコア技術の開発だけで、システム開発はやっていません。何かプロジェクトをするときはSIerと組むことが多いですね。

**——なるほど。SIerは各業界の顧客のニーズが分かっていますからね。**

▶ 中期的には自動運転、ヘルスケアの領域も

**——では、もう少し先にはどのようなことを手がけたいと思っていますか？**

中期的には、自動運転とヘルスケアの領域に重点的に取り組もうとしています。

自動運転のコアな部分は、トヨタと共同で研究開発を進めています。

秘密保持契約があるので多くを語れませんが、個人的には、今の契約の先、つまり自動で車を制御できるようになったあとは、車間ネットワークを通じて交通全体を最適化することを目指したいです。道路インフラと連動してうまく車を誘導し、都市全体で輸送を最適化する。その辺まで行きたいと思っています。

実は自動運転車が人間の運転能力を超えるのは、まだまだ先だと思っています。人間には高度な予測能力があるからです。

ただ機械のよさの1つに、一瞬で同期できるということがあります。

1台の車の運転では人間に勝てなくても、車と環境のネットワークがインタラクションし、人間が見えない情報を取り扱うことによって、人間の運転能力を越えることも可能になると思います。例えば車同士や街頭のカメラと連携して、1台の車からは見えない先々の渋滞情報や周りの状況が分かるようになって、事故を回避できるようになるのだと思います。

単体のデバイスの頭をよくすることは、Googleなりロボットの会社がやっていくでしょう。われわれはそこは狙わないで、デバイスが結びつくことによって初めて可能になるような価値を作っていきたいと思っています。

## ——ヘルスケアに関してはどのように取り組んでいくのですか？

ゲノムを解析してどのような疾患にかかりやすいかを予測するサービスは既に出てきてますよね。でもIoTが発展することで、人の状態を常に取り続けることになると思います。単に生体データを取るだけでなく、ゲノムの情報も毎日トラッキングできるようになると思います。

## ——遺伝子解析は毎日するようになるんですか？

はい、そうなると思います。遺伝子解析の機械ってシーケンサーと呼ばれて、今はまだ高額で1000万円くらいします。でも小型化が進んでいて、5年後にはより低価格になり、USBメモリサイズのシーケンサーが普及し始めるとわれわれは考えています。

そうすると空気中のウイルスや最近などのDNAの情報を、常にトラッキングできるようになります。インフルエンザのワクチンを作るには、今年は何のインフルエンザウイルスが流行ってきているのかが分からないといけない。そのためには人口の何パーセントかにまずインフルエンザにかかってもらう必要がある。ところがシーケンサーが低額になると、空気中の細菌やウイルスのDNAを解析できるようになり、どういうインフルエンザウイルスが増えてきているのか正確に把握できるようになります。

風邪で病院に行ってもウイルス性か細菌性か分からないので、とりあえず抗生物質をもらうということもよくある話です。でも喉のところのシーケンシングができれば、ウイルス性か細菌性か、どういうウイルス

か細菌かが正確に分かる。そうなれば不必要な薬を投与することがなくなります。

**——ウイルスや細菌を特定するためのシーケンシングというわけですね。**

そういう使い方が1つ。もう1つは自分のDNAを日々トラッキングできます。

**——日々トラッキングして、どのようなメリットがあるのですか？**

例えば癌はDNAの変化なので、日々血流のDNAをトラッキングすることで早期発見が可能になります。今でも問題の遺伝子を特定できていれば、その遺伝子を解析して変化を比べることは可能です。でもまだ問題が明らかになっていないときには、高額過ぎて毎日シーケンシングするわけにはいきません。

でもだれでも手軽に安価に自由にシーケンシングができるようになると、日々シーケンシングするようになり、病気を早期発見できるようになります。基礎技術が5年後ぐらいにはできるでしょうから、医療の世界が大きく変わるのではないかって思っています。

**——そこに機械学習はどう関与してくるのですか？**

まずはシーケンシングの実験そのものに機械学習が入ってきます。シーケンシングにはミスが発生することが多いんです。シーケンシングは人が行いますから、人によって実験環境が変わります。培養するときの温度が違えば、違った結果になったりします。実験が得意な人、不得意な人がいます。

なのでわたしは実験は機械化すべきだと思っています。どのような温度、どのような環境なら、どのような結果になるのか。それを機械で学習して、適応させていく。そういうカタチでシーケンシングに機械学習を入れていきたいと思っています。

さらに、診断そのものに機械学習が入っていくと考えています。今はこういう遺伝子があるから、この病気になりやすい、という程度の判断

しかできていない。ただ実際には生命の仕組みはもっと複雑なわけ  
です。いろんなDNAがいろいろ絡み合い、いろんなタンパク質が生成さ  
れ、それらが相互に作用し、いろんな要素が組み合わさって病気にな  
る。非常に複雑なプロセスを経て病気になるわけです。その複雑な仕  
組みを機械学習で解明していきたい。

生命は、複雑なネットワークで動いている。どんなインタラクションが  
行われているのか、複雑過ぎて網羅的に解析することができないんで  
す。いろんな人種がいて、いろんな遺伝子のパターンがあり、どういふ  
うにタンパク質がインタラクションしているのかを解釈するのは、とて  
も人手による実験ではできない。人の脳で理解するには複雑過ぎて不  
可能に近い。

なのでわれわれは人の中で起こっているデータを全部収集し、それ  
を機械によって解析させる。そういうところをやろうと思っています。そ  
れには、機械学習を始めとするパターン認識のアルゴリズムが重要に  
なってくる。

既に副作用の解析では、研究レベルで機械学習が使われ始めて  
います。

言語解析の例でも、実は深い階層にはラベル付けされていない重  
要な情報のインタラクションがあるのではないかという話をしましたが、  
生命の仕組みも同様のことが言えるんじゃないかと思います。もし本  
当にそうであるならば複雑なパターンを認識する仕組みが必要。複  
雑なインタラクションのパターンを学習できるような機械学習のアルゴ  
リズムが必要になってくる。

そこでDeep Learningの機械学習が使えるのではないかとわれわ  
れは思っているわけです。

実はUSのメンバーの一人に、ずっと創薬分野の研究をしていた人  
間がいます。シミュレーションベースで化合物のカチを予測して創薬  
に結びつける。そんなコンピューターシミュレーションの研究を続けて  
きたんです。彼が今、アメリカの子会社のCOOをやってくれています。  
彼が中心になって、Deep Learningを使った薬の副作用を解析して  
みようとしています。人によって副作用が起こりやすいのか、起こりに  
くいのかを、コンピューターを使って解析しようという試みです。研究レ  
ベルではバイオの世界で行われ始めていて、恐らく5年後ぐらいには、  
その分野の研究はかなり進むのではないかって思っています。

——そうした自動運転やヘルスケアなどの中期的ビジョンの実現に  
は、新興国のほうが有利だったりしませんか？

そうですね。そうだと思います。ヘルスケア、スマートシティ、自動運転などは、新興国のほうが先に進む可能性があります。特に自動運転、スマートシティなどは、都市を一から作る段階から導入するほうが、面倒な承認手続きなしで進められるからです。

そういう意味でシスコは、われわれにとって理想的なパートナーだと思います。彼らは、世界規模でそういうところとの関係を持っている。

彼らのデバイスにわれわれの頭脳を搭載することになれば、彼らは瞬時に世界中の彼らのデバイスにプログラムをデリバリーすることが可能になります。われわれが営業活動する必要がないわけです。

**——シスコのデバイスってパソコンのルーターのようなイメージしか持っていないんですが、IoT的にはどのようなデバイスを出しているのですか？**

シスコは今まで、ネットワークの中心部分、コアのネットワーク機器を主に製造していたのですが、最近は10万円から数十万円のネットワーク・エッジ向けのデバイスを手がけています。自動車などに搭載可能なデバイスです。

中でもIOXと呼ばれるルーターは、ルーターとリナックスマシンが一体化しているのが特徴です。またそれだけではなく、ストリームの流れてくるデータを必要に応じてフィルタリングして上位にアップロードしたりできます。また必要なときに必要なデータをうまく取り出せるようなAPIが備わっているんです。

最近のシスコは、単純なパケットのルーティングだけではなく、ルーターの中で、できるだけインテリジェントな処理ができるプラットフォームを提供しています。例えば、工場でたくさんのセンサーデバイスをシスコデバイスにつなげれば、得られた情報を元に、例えば「このレーンを停止せよ」とかのイベント処理まで、ネットワークの一番端でやってしまいます。

ただ今のところそういう仕組みは、ルールベースの仕組みなんです。どういうデータが来れば、どういうアクションを起こすべきか、というようなルールを人が記述しています。実は、そのルールの記述が今はボトルネックになりつつある。決められたルールでは、不測の事態に対応し切れないからです。なので不測の事態に対応できるような仕組みを、われわれの機械学習の仕組みで実現しようとしています。

——シスコはスマートシティではどのようなことをやっているのですか？

例えば信号機のところにカメラがついていれば、横断歩道を渡りきれない老人の認識が可能です。人を認識すれば、歩行者用信号の時間を延長することが可能になります。

——スマートシティになれば、中央分離帯が交通量によって移動したりするようになるのでしょうか？

そんなことしなくても、自動運転が可能になれば、中央分離帯も白線も不要になると思いますね。

▶ 宇宙のなぞの解明にも

——人間の手に負えないような複雑な仕組みの解明に機械学習が有効だということですね。そうであれば、宇宙の謎の解明には使えないのでしょうか？

使えます。実は、国の宇宙のプロジェクトのお手伝いをしようとしています。われわれが提唱するエッジ・ヘビー・コンピューティングって、クラウドの中央コンピューターにデータを集めるのではなく、もっとエッジの部分でもっとインテリジェントな処理をさせる、という考え方です。シスコのフォグ・コンピューティングも同じような考え方です。そうした考え方が、宇宙の謎の解明に非常に有効だと思います。

宇宙にはたくさんのデータが飛び交っています。ただ、今はそれを全部地球に持ってきて解析しようとしている。例えば電波望遠鏡などでつかまえて解析しようとしているわけです。また人工衛星が、地球の情報を始め、宇宙のたくさんの情報を集めることができるけど、地上にはあまり転送できないんです。

大気が邪魔をして、通信速度が著しく落ちてしまうんです。雲が発生するだけで通信速度が激減してしまう。宇宙のデータを解釈しようとしても、今は十分なデータを集められないんです。一部サンプリング



された情報しか集められていないのが現状です。

でも宇宙ステーションが機械学習の仕組みを持ってしまえば、地上にデータ転送しなくてもよくなる。宇宙に人工衛星をいっぱい飛ばしておけば、宇宙での人工衛星間の通信は光通信でもものすごく速い。それを宇宙で処理してしまえば、大気圏の通信速度の減退の影響を受けない。データの解像度を落とす必要がない。まずはそうした仕組みを作ろうとしています。

**——宇宙の謎にも取り組んでいるのですか。すごい話ですね。ところでより先の未来はどういうものを目指していますか？**

長期的には、ネットワークアーキテクチャーが全然変わっていくと思っています。

今のネットワークは、基本的にはデータを通すためだけのネットワークなんです。われわれは、それを人間の神経系のようなものにしていきたいと考えています。人間の神経系は、情報が脳に行くまでにいろんなフィルターを通過させて、適切な情報だけが脳にくるようになっている。判断に必要な情報だけがサマライズされて脳に集まってくるんです。そうした神経の高度な仕組みをネットワークに適用していきたいんです。

自然言語処理の仕組みとIoTが融合して、言語の理解が進むという話をしましたが、いずれインターネットとニューラルネットって融合していくのではないかと思います。世界全体がインターネットを介して高度に分散されたデジタルな神経系になっていく。そういう世界を目指しています。

その仕組みに当然、人間をアタッチすることもできます。ブレイン・マシン・インターフェースのようなものが今後、発展していくと、見えないところのものも情報が見えるようになるかもしれない。高速道路を走っているときに、この先の渋滞箇所の映像をみたいと思えば、スクリーンやデバイスを参照することなく、映像を見ることができるようになるかもしれない。つまり人間の神経系とデジタルな神経系が将来的には、くっついていくようになるのではないかと思うわけです。

**——脳に電極を埋め込むようになるのでしょうか？**

非接触で、できるようにならないかなって思ってます。人間のデータ

を吸収する仕組みや脳の仕組みがもうちょっと分かれば、電極を刺さなくても人間の神経系に影響を与えることができるようになるかもしれない。そうなればだれもが仮想的にインターネットにつながるようになるように思います。

▶IoTデバイスは変幻自在に

——究極のIoTネットワークって、どのようなものになるのでしょうか？

今のIoTデバイスってデータを集めるのが主な目的ですよ。集まったデータは中央で処理される。デバイス自身が環境に対して最適化されているかという、そうではないです。デバイスはカタチが決まっています、そこからデータを取るだけ。そうではなく、デバイス自身が解きたい問題に対してカタチを変えていくようになるのではないかと思います。

というのは人間は、触感で物のカタチを認識するときに、指を曲げて、物のカタチに合わせてるじゃないあですか。それと同じことがIoTデバイスでもできるようになるんじゃないかって思います。

——指の先だけで触れて、その触感だけでモノのカタチを判別するのは難しいけど、カタチに合わせて指を曲げて手全体で感じることで、より簡単にカタチを認識できる。そんなふうにIoTもなっていくわけですか？

そうです。頭脳部分とリアルタイムでインタラクションしてデバイス自身のカタチが変わっていく。これからのデバイスはそうなるべきだと考えています。環境に対して適応してデバイス自身が自動的にコンフィグレーションしてカタチを変え、来た問題に対してエネルギーを最適化するように動いていく。そうしたデバイスを将来的に作っていきたい。

そのために重要になるのが、デバイスと脳をつなぐための神経の部分。それを作っていないといけない。そういう意味でネットワークはまだまだ進化する余地があると思っています。最終的には、われわれはコンピューターネットワークの会社になっていくのだと思います。

## ——今のIoTネットワークと、理想的なネットワークの違いをもう一度説明してください。

今のIoTネットワークは何も解釈せずに一方向にデータを通すだけ。今後のネットワーク・アーキテクチャは、常に双方向で、かつ得られた情報を瞬時に解釈しフィードバックしていく。自律制御システムのようなものになると思います。

単体のデバイスの中では自律制御はできているけれど、ネットワークを介してしまうと、今は途端にできなくなる。われわれは、ネットワークを通じてリアルタイムにフィードバックを送れる仕組みを作ろうと考えています。知識がその場で生成され、知識を通す仕組みになるわけです。つまり知らない情報をローカルでフィルタリングして、ネットワーク自身でもフィルタリングする。

例えばカメラが3台違う場所から同じものを写していたとします。重なっている撮影領域があるので、重複するデータがあります。それをすべてを中央に送るのは無駄。といっても個々のカメラからは、どれがわからない情報か分からない。でも中央に送る前に途中のルーターでそれぞれのデータを照合し、要らない情報を判断することができます。そして必要なデータだけを中央に送るようになります。そこより上へのデータ転送量が少なくて済むわけです。

一方で3台が同じものを撮影しているのが無駄なので、カメラの向きを調整するように、ルーターがカメラに命令することもできます。つまりネットワークデバイスが効率的に判断するわけです。それが今までのネットワークデバイスとの違い。今までのネットワークデバイスはセンサー制御までしてこなかったんです。

## ——ネットワーク全体が「考える」わけですね。

そうです。生物と同じです。生物は、感覚器がたくさんついていて、いろんな環境とインタラクションするからこそ、脳が成長していく。脳だけ取り出して成長するかというと、そんなことはないです。

たくさんの情報を効率的に収集して学習できるようにしているのは、脳だけではないんです。脳の学習能力は非常に強力なんですけど、データをどうやって効率よく集めるのかは全身で行っている。

一方で、今のクラウドコンピューティングは脳に全部集めましょうという考え方。集める仕組み自体については、あんまり考えていない。

僕らが目指しているのは、データの集め方も変化するようなネットワーク。脳が最も賢くなるように、集め方を工夫するネットワーク。単にデータを集めるだけではなく、もっとうまく集められるようにデバイス自身も変化するネットワーク。そういうネットワークを世界に広げたい。それがわれわれが目指している世界です。

**——自動車にしる、ヘルスケアにしる、大手が強い領域のように思っています。ベンチャーも大手と組むのが有利な時代になってきているのでしょうか？ スマホ時代のように、だれもがアプリを作れてスタートアップが乱立するという時代ではなくなったのでしょうか？**

そうですね。これからは、デバイスへの理解が必要になると思います。デバイスにはアナログの要素が含まれます。自動車なんかは特にそうですね。自動車の作り方を理解していないIT系の人には、絶対にいいものを作れない。

これからいろいろなデバイスにITを組み込んでいくには、ITの人はそのアナログの部分の理解しなければなりません。Googleが自動車メーカーから大量に人材を採用しているのは、そのためです。

ヘルスケアも同じだと思います。われわれも生物学をしっかりと勉強しないといけないと、メンバーの間でいつも話しています。今は、京都大学のIPS細胞の研究チームと組んで研究を進めています。

コンピューターサイエンスだけでは、IoTを発展させることはできないんです。なのでIT産業の中心が移っていくのかなとも思っています。

ウェブの世界を離れてIoTの世界になると、従来型産業の大手企業が強い部分もまた出てくる。でも大手がコンピューターサイエンスを正しく理解しているかという、そうではない部分も多い。両方を理解して研究を進めていかないと、IoTは発展していかないんじゃないでしょうか。

**——起業家は、どこに努力とリソースを集中すればいいのでしょうか？ これからはどこにビジネスチャンスがあると思いますか？**

われわれがネットワークを整備し、人間の神経系のようなインテリジェントなネットワークを世界に広げていきます。われわれがインフラを作るわけです。それを使ってどのようなビジネスができるのか。今までにはない環境が可能になるわけですから、ビジネスチャンスはいっぱいあると思います。

将来的には、ネットワークデバイス向けのアプリストアみたいな仕組みを提供していきたいですね。

**——スマホアプリ市場を通じて、新たなスタートアップが無数誕生したように、IoTアプリ市場ができれば新たなビジネスチャンスが生まれそうですね。**

そうですね。そうした世界の実現のために今理想的なパートナーは、シスコ。IoTはいろんな会社に取り組んでいますが、いろんな会社と話をする中で、シスコの取り組み方に一番将来性があるように思います。

シスコがおもしろいのは、ネットワークの会社でありながら、機械学習の専門家が率いているチームがあるということ。ネットワークに機械学習を取り込んでいこうということに対してビジョンを持っている。それにびっくりしました。

西川氏とのインタビューの中で、シスコ・システムズのフォグ・コンピューティングのことは以前メルマガ向けの記事にしたことがある。参考までにここに掲載しよう。

## クラウド（雲）からフォグ（霧）へ FOG COMPUTINGが開くIT新時代

クラウドコンピューティングという表現がIT業界関係者の間で一般的に使われるようになったかと思えば、今度はクラウド（雲）ならぬフォグ（霧）という言葉を含むFog Computingという新しいキーワードが米国のIT業界で使われ始めた。

クラウドコンピューティングは、自社のサーバーを使ってデータを記憶させたり演算処理をさせたりするのではなく、インターネット上の専門業者のサーバーを使って同様の業務を行い、その使用料だけを支払うという事業の概念。どこにどのような形でサーバーが設置され、どのように使われているのか分からない。まるで雲の上のサーバーを利用してのような感じがするので、クラウド（雲）コンピューティングという言葉が使われるようになった。

クラウドコンピューティングは、サーバーの初期投資が不要なので大きな資金がなくてもベンチャー企業を起業しやすくなったし、大企業にとっても効率よく運用できるメリットがある。なので、クラウドコンピューティングは利用が進む一方だ。ただ運営コストを下げるために、土地や電気代が安い場所にサーバーを設置するため、通信にわずかながらも時間がかかる場合がある。私自身の経験でも、スマートフォンから航空券の予約をしようとすると、空席を表示するのに数秒かかり、ちょっとイライラしたことが何度かあった。

一方フォグコンピューティングは、使った分だけ使用料を支払うという事業形態はクラウドと同じだが、顧客の特定用途の端末の近くにサーバーを設置することで、通信にかかる時間を極力小さくすることができるのが最大の特徴。霧も、雲のように水蒸気で視界が遮られ、中で何が行われているのか分からないが、霧は雲よりもより近くに存在する。そういう意味でフォグ（霧）コンピューティングという言葉が使われ始めている。

同様の概念は、これまでエッジコンピューティングなどと呼ばれることもあった。しかしスマートフォンを中心とした無線通信が主流になったり、あらゆるモノにセンサーと通信機能が搭載されるInternet of Things（IoT、モノのインターネット）の時代が本格的に始まる中、パソコン時代のエッジコンピューティングとは異なる仕組みが必要になってきた。なのでフォグコンピューティングという新しい名称で、そのビジネスの可能性を考えなおそうという機運が高まっているようだ。

この名称を提案しているのはルーター大手の米シスコ・システムズだが、米ウォール・ストリート・ジャーナルなども最近の記事で、この名称を取り上げて、「フォグがテックの未来だ」と論評するなど、この名称が広く受け入れられるようになりそうな雲行きだ。

## ★フォグコンピューティングの7つの特徴

シスコによるとフォグコンピューティングの特徴は次の7つ。1つ目は、通信に時間がほとんどかからないことと、場所が重要な要素であること。「土地代、電気代が安ければ地球上どこにサーバーを置いても構わない」というクラウドコンピューティングとは異なり、フォグコンピューティングではよりリアルタイムにスムーズにデータを転送するために、サーバーの設置場所が重要になる。ゲームや動画のストリーミング、拡張現実（AR）などのサービスは、通信に時間をかけずに大量のデータを送受信しなければならない。これらのサービス向けにネットワークのあり方を工夫したことが、フォグコンピューティングの始まりだという。

この延長線上で今後、自動車の後部座席向けの映画配信サービスなどのビジネスが登場するかもしれない。そのサービスを実現するために、高速道路や幹線道路沿いにプロキシサーバーを設置する事業者も出てくることだろう。今後、どこにサーバーを設置するかがますます重要になっていくことだろう。

2つ目は、端末が広く点在していること。スマートグリッドは電力の流れをセンサーがモニターし、最も効率よく電力を配分できるようにする仕組みだが、電力網上のセンサーは広範囲にわたって設置されることになる。この広く点在する端末からのデータを、インターネット上で他のデータの交通の中に混ぜて渋滞を引き起こすのは、あまりに非効率。電力網上のセンサーのデータだけを集計する仕組みを構築すれば、効率よくリアルタイムに電力の流れを自動制御できるようになる。スマートグリッドは、フォグコンピューティングの代表的な用途になるという。

3つ目は、端末が移動することがあるということ。先の自動車の後部座席向け映画配信サービスなどがその典型例だが、移動し続ける端末に対して、途切れることなくデータを送り続けることのできる仕組みが求められるようになる。

4つ目は、端末数がかなり多いこと。IoT時代には、ありとあらゆるモノにセンサーが搭載される。その数の多さに対応できるシステムである必要がある。

5つ目は、ワイヤレス通信が中心となること。このことは、エッジコンピューティングと呼ばれていた時代には、それほど重要ではなかった。しかしフォグコンピューティングのシステムは基本的に、ワイヤレス機器に効率よく対応することが最重要課題の1つになる。

6つ目は、ストリーミングやリアルタイム性が重要なアプリケーションであること。クラウド型ではデータの送受信を効率化するために、データがある一定量集めてから一括処理する「バッチ処理」と呼ばれる手法を取ることが多い。これに対し、ストリーミングやリアルタイム処理が必要なサービス向けに、サーバーの設置方法に工夫を加えたフォグコンピューティングが今後増えてくることだろう。

7つ目は、多様性に対応できること。サーバーにアクセスしてくるのがWindowsパソコンが中心だった時代は終わった。スマートフォンに加えて、各種ウェアラブル機器、センサー機器が一斉にアクセスしてくるようになる。それに対応できる仕組みが不可欠になる。

#### ★まずは自動車、電力、ヘルスケアで

では、フォグコンピューティングはどのような領域で利用されるようになるのだろう。シスコが2012年にまとめたレポートでは、コネクテッドビークル（IT化された自動車）、スマートグリッド、ワイヤレスセンサー、スマートシティ（IT化された自治体）、ヘルスケアなどの領域に特化したクラウドサービス、つまりフォグコンピューティングが台頭してくるだろうと予測している。

シスコが予測するフォグコンピューティングの用途の1つとして、スマート信号システムが興味深いので紹介しよう。

スマート信号は、信号付近の自転車や歩行者を察知すると同時に、遠くから近づいてくる自動車の距離や速度から信号まで到達する時間を瞬時に計算。その上で周辺の自転車、歩行者、自動車とって何色の信号を何秒間表示するのが全体にとっての最適解になるのかを計算し、その信号の色を表示する。からかじめ色を変える時間が決まっているため、歩行者がいなくてもかわらず、赤信号で自動



車を止めるようなことがなくなるわけだ。交通がよりスムーズに流れることになる。

周辺の信号機とも連携をして青信号の表示を続けることで自動車の速度を高めたり、反対に制限速度を超えて走る自動車の速度を落とすために赤信号表示を増やすことも可能。信号システムのほうで自動車の動きをコントロールできるわけだ。

ではどういう企業がフォグコンピューティングのサービスを提供するようになるのだろうか。クラウドコンピューティングの領域で、Amazonが自社のサーバー運営のノウハウを利用して他社向けにクラウドコンピューティングサービスを提供しているように、フォグコンピューティングの領域でもユーザーとなり得る事業者がサービス提供者側に回ることも考えられる。

シスコのレポートによると、利用者側、提供者側のどちらになるかは分からないが、電力会社、ガス会社、電話会社、自動車メーカー、自治体、鉄道会社などもフォグコンピューティングのプレーヤーになる可能性があるとしている。

Wall Street Journalの記事

[Forget 'the Cloud'; 'the Fog' Is Tech's Future](#)

シスコシステムズのレポート

Fog Computing and Its Role in the Internet of Things



## 第2章 協働型ロボットの今

湯川塾24期第2回

日時：2014年11月10日（月） 19:30～21:30

場所：渋谷co-ba

講師：影木准子氏

北海道大学工学部を卒業後、日本経済新聞社で13年間、記者として働く。うち1997-2001年の4年間は同社シリコンバレー支局勤務。シリコンバレー在住のフリーランス・ジャーナリストを経て、現在はカワダロボティクス社企画部課長。世界で活躍する女性ロボット関係者25人の1人にも選ばれている。

影木さんは、日本経済新聞の元記者。実は影木さんがシリコンバレー支局特派員時代の1990年代後半に、僕は時事通信社のサンフランシスコ支局（現シリコンバレー支局）にいた。なので、取材先でお仕事をご一緒したこともあるし、プライベートでも親しくお付き合いさせていただいていた。

影木さんはその後、シリコンバレー勤務を終えて2001年に帰国。シリコンバレーでお付き合いをされていたドイツ人ジャーナリストの方と結婚されて日本経済新聞社を退社し、シリコンバレーに戻られた。シリコンバレーではフリーのジャーナリストとして、ロボット専門のブログメディア&メルマガ「GetRobo」を立ち上げられた。

このブログメディアは、業界内で高い評価を得ていたようで、24期に参加された塾生の一人も「GetRobo」の読者だった。ただ、発表するメディアの激減、原稿料の減少など、フリーのジャーナリストにとっては厳しい状況が続いたため、転職を決意し、メルマガの読者だった川田工業株式会社に就職することになった。（後に川田工業が子会社のカワダロボティクス株式会社を設立し、そちらに転籍。）

「カワダロボティクスを選んだ理由は、たくさんのロボットを見てきた中でこのNEXTAGEが最も好きだったからです。コミュニケーションロボットという領域もありますが、わたしは作業をするロボットが好きなんです」と影木さんは言う。

影木さんによると、世界のロボット業界では今、コラボラティブロボットという領域がホットになりつつあるのだという。略称コボット、日本語では、協働型ロボットと呼ばれている。

人間と一緒に作業ができることが特徴だ。従来の産業用ロボットは、人間と隔離されて設置されていた。腕を振り回すロボットの近くにいると、振り回してきた腕にぶつかるのかなり痛いからだ。「痛いどころか、死んでしまう可能性だってあるんです」。

そうならないように、安全性を考えて作られているのが、協働型ロボットだ。協働型ロボットは、隔離する必要がなく、人と同じ環境で働かせることが可能だ。「労働安全基準法では、80ワット以下のモーターで動く機械は産業用ロボットに該当しないので、法的には困う必要がないんです」と影木さんは言う。ただロボットを購入した企業が、自分でリスクアセスメントして、黄色のテープで囲ったりすることはあるようだ。

カワダのロボットは、モーターを15個搭載しているが、すべて80ワット以下で駆動する。ただ2013年には法改正があり、80ワット以上でも安全性が確保されればロボットを囲う必要がなくなったのだという。

## ロボット台頭の背景に労働者不足

カワダのロボットは、人の形をしている。「ヒューマノイド・ロボットにこだわっている」のだという。どうして人の形にこだわるのだろうか。

「人の置換えではなく、一緒に働くことを目指しているからです。人と作業を交代できることを目指しているんです」と影木さんは言う。

パートさんのお子さんが風邪を引いて、急に休まないといけなくなった。来週、急に増産が必要になったけど、今からパートを募集できない。そういう状況に対応するためのロボットなのだという。ということは、今までパートさんが立っていた場所に入ってもらわないといけない。なので背格好が人間と同じに作ってあるのだという。

日本経済新聞によると2014年は大労働力不足時代元年なのだとか。工場が街から離れた場所に設置され、パートタイマーがあつまらない。たとえ集まって3カ月トレーニングしても、飽きて辞めてしまう人が多い。この3ヶ月トレーニングのコストがバカにならない。こうした状況が、協働型ロボットが求められる背景になっているようだ。

確かに、人の形をしていることで、人間の生活環境、職場環境を変えずにそのまま利用できる。われわれの社会環境は、すべて人間の背格好に最適化されている。ドアノブは人間の手が回しやすい大きさに作ってある。階段は、人間の歩幅にちょうどいい高さに作ってある。なので、ロボットを人間の背格好と同じくらいにすれば、今の社会環境を変えることなく、そのままロボットに活躍してもらえることになる。

そしてせつかく人間の環境に合わせてロボットを人間に似せるのであれば、人間の作業のほとんどをロボットにしてもらいたいもの。工場の大型産業用ロボットが専門の作業に特化されているのに比べ、人型ロボットには汎用性が求められるのはこのためだ。

川田工業のロボットNEXTAGEも、汎用性が高く設計されている。ロボットもコンピューターの一つなのでプログラムで動いている。どう動かすかは、このプログラム次第なのだが、カワダロボティクスではプログラム作成を「プログラミング」とは呼ばずに「ティーチング」と呼ぶのだという。プログラミングスキルのない人でも、表計算シートにデータを書き込むような単純な作業で、NEXTAGEをプログラミングできるからだそう。

## 飲食店向けロボットは実現するのか

この汎用性の高さが、ロボットの普及に関して非常に重要なことなのだと僕は考えている。

というのは、コンピューターの世界でも、コンピューターが爆発的に普及するようになったのが、「低価格化」と「汎用性」の2つの要因が合わさったときだったからだ。30年以上前、コンピューターといえば、専用の作業しかこないせない大型コンピューターが中心だった。ところが、プログラムを書き換えることができるコンピューターが登場した。メインフレームと呼ばれるタイプの大型機だ。メインフレームの日本語訳は、「汎用機」だった。

そして1台30万円という低価格コンピューターが登場した。パソコンだ。「汎用性」と「低価格」の両方の要素が合わさった。そのころからパソコンは急速に普及し始め、そしてついにはオフィスのすべてのデスク上に1台ずつ置かれるようになった。

NEXTAGEに見られるように、ロボットの汎用性が高まりを見せている。そしてソフトバンクのロボットPepperは20万円という低価格を実現した。「汎用性」「低価格」という2つの要素を併せ持ってパソコンが急速に普及したように、ロボットがこれから急速に普及し始めるかもしれない。そう思うわけだ。

カワダロボティクス社のNEXTAGEはどれくらい汎用性があるのだろうか。飲食店などでも使えるのだろうか。

「飲食の現場にはまだ入っていません」と影木さん。その理由は2つ。1つは、スピードが出ないからだ。「周辺環境が変化する場合、ロボットは絶対人よりは速くならない」という。まずは画像の認識に時間がかかるからだ。

また人間は万能な一組の手であらゆることをやってのけるのだが、ロボットは取り扱うものによって「手」を変えないといけないのだと言う。部品をつかむ時の手、蓋を閉める時の手が別の手というものがある。人と同じ周辺装置と道具を使う事が出来るのだが、使う道具によってロボットの手を交換する必要がある。もちろん手は自動で交換できるのだが、交換のために時間がかかる。その結果、同じ作業時間で人間なら10個の製品を作れるところを、ロボットだと5個から9個しか作れない。

もう1つの理由は、食品を取り扱う際に要求される清潔さをどう確保するか。洋服を着せるのがいいのか。精密部品の固まりであるロボットを洗浄するのは困難だ。

画像認識技術は、人工知能の急速な進化で今後一気に改良されるのではないかと見られている。一方で、手の交換の問題が、どれだけ大きな課題なのだろうか。解決策があるのか。僕には分からない。

手の交換の問題が解決されたときに、ファーストフードチェーンのアルバイトの仕事がロボットに取って代わられるようになるのか。「時々刻々と変化する環境下では人よりも絶対に速くならない」ので、やはり飲食店にロボットが普及することはないのか。

「ロボットがあらゆる仕事を奪っていく」と無責任に予言するのは簡単だが、汎用性ロボットの進化を、もう少しウォッチしておく必要があるようだ。

カワダロボティクス社の川田グループは、もともとは建設会社としてスタートしている。なので建設現場で利用できるロボットを開発することが、もともとの目標だ。カワダロボティクスが目指す建設現場向けロボットの開発の動向も気になるところだ。

## ロボットはすべての仕事を奪うのか

米Northwestern大学経済学部のRobert Gordon教授は[米BBC放送のインタビュー](#)に対し、ロボットがあらゆる仕事を奪っていくという意見に対し懐疑的な見解を示している。

同教授は、毎日出会う人を眺めては「この人の仕事は、ロボットに取って代わられるんだろうか」と考えるようにしているのだという。しかし「すぐにロボットに取って代わられそうな仕事はほとんどない」と言う。例えば、宅配便の運転手は、ただ運転することだけが仕事ではない。配達先の住所に着くと、運転席から降りて、後ろに周り、荷物を荷台から下ろす。それから荷物を持って階段を登って運ぶ。「これだけのマルチタスクをこなせるロボットは今のところいない。そういう多機能ロボットが登場するまで、まだまだ時間がかかる」と言う。「現在のロボットは1つの作業に特化したものばかり。長期的未来には多機能ロボットが、一部人間の仕事に取って代わる可能性はあると思うが、まだ遠い先の話」と語っている。

### ▶脚光を浴びるBaxter

「汎用性」と「低価格」という要素を兼ね備えたロボットとして世界的に有名なのが、[rethinkrobotics社](#)のbaxterだ。

影木さんによると、同社が注目されている1つの理由は、お掃除ロボット「ルンバ」の開発者であるロドニー・ブルックス氏が会長兼CTOを務めているかららしい。また、AmazonのCEOのジェフ・ベゾス氏が個人的に出資していることでも有名だという。

またbaxterの最大の特徴はダイレクトティーチングと呼ばれるプログラミング手法で、難解なプログラミング言語を使用する必要もないし、キーボードでデータを入力する必要もない。人間が直接ロボットの腕を動かして、動作を覚えさせることができる手法。

ただ、簡単に教えられるという利点がある一方で、動作の精度はよくないようだ。というのは、このロボットは安全性がウリの1つで、人や物にぶつかった際に動作を無理に継続しない設計になっている。業界用語で言うところの「負ける」仕組みになっているというわけだ。ぶつかったら「負ける」ように設計されているので、どうしても精度がでないようだ。日本でフル装備430万円。かなり安くなってきているが、

# 急展開する業界状況

今はカワダロボティクス社で、バリバリ働いている影木さんだが、もともとは経済ジャーナリスト。しかもフリーになってからは、シリコンバレーでロボットを専門にしてブログメディアを運営していたのだから、業界動向、特にシリコンバレー周辺の動向に関してはだれよりも詳しいはず。

そこで影木さんから見たロボット業界の最新の動向で面白そうなものを幾つかピックアップしてもらった。

## ①Amazonの動き

仮想本屋最大手のAmazonが精力的に動いている。2012年にはロボットベンチャーのキバ・システムズ社を買収している。キバ・システムズ社が開発したのは、棚が移動する仕組み。これまで、人が棚まで歩いていたのだが、棚の方から人のところまでやって来るという仕組みだ。

Amazonは別にキバ社を買収しなくても、キバ社からこの移動式の棚のシステムを購入するだけでよかったはず。恐らくキバ社を買収することで、競合のEC事業者はこのシステムを使わせたくなかったのかもしれない。

またAmazonは[ピッキングチャレンジ](#)と呼ばれる開発コンテストを主催している。倉庫内でのピッキング（集荷）作業のロボット化に関する技術開発コンテストで、参加者にはイベント参加の経費や、必要な機材を提供するほか、優勝者には豪華賞品を提供している。ピッキングは、倉庫、流通業の次にキー・テクノロジーと言われている。このコンテストを主催することで、次世代技術を競合他社に与えないという戦略ではないかと思われる。

「・・・チャレンジ」と呼ばれる開発コンテストは、もともとDARPA（アメリカ国防高等研究計画局）が行うダーパ・チャレンジから始まったもので、最近米国では同様の「・・・チャレンジ」が頻繁に開催されているという。ロボットの学会は、春のICRA、秋のIROSがあり、どちらかに参加すると最先端の情報を得ることができる。Amazon・ピッキング・チャレンジは2015のICRAで行われる予定。

Amazon・ピッキング・チャレンジでの推奨プラットフォームとなっているロボットは、[バクスター](#)、[ユニバーサルロボット](#)、[PR2](#)のなど5つ。特にPR2の開発元であるウイローガレージ社はロボット業界を変えたと言われている。同社はロボットオペレーティング・システム



ROS（OSはリナックス、ROSはミドルウェアみたいなもの）を開発、ROSと統合された「PR2」11台を世界の有力開発者に無償で配った。これでROSはロボット業界のデファクト・スタンダード（事実上の業界標準）になった。トヨタのHSRもROS対応。「トヨタがオープンソースを使うのは驚きだった」と影木さんは語っている。

②従来の産業用ロボットメーカーも協働型ロボットに参入してきている。

- ・ABBはFRIDA発表 まだコンセプトモデル
- ・ファナック
- ・安川電機
- ・セイコーエプソン

③2013年Googleは、ロボット関連の会社立て続けに買収した。

- ・ボストンダイナミクス：油圧駆動ロボット
- ・シャフト：電動駆動ロボット。東京大学情報工学研究所のスピアウト。二足歩行ロボットを研究
- ・メカ・ロボティクス：ヒューマノイド・ロボットを、主に研究分野で利用
- ・レッドウッド・ロボティクス：低価格のロボットアーム
- ・インダストリアル・パーセプション：コンピュータービジョン
- ・ボット&ドリー：ロボットを商業的に使うのがうまい
- ・ホロムニ：可動式車輪の開発
- ・オートファス 広告代理店
- ・ネスト・ラボ：サーモスタット

しかし、責任者のアンディー・ルービン氏が辞めてしまった。これはどういうことなのだろうか。Googleはどんなロボットを開発しているのだろうか。別の業界筋は「これは単なる憶測ですが」と断った上で、「ロボットベンチャー企業は、どこも天才的研究者がトップ。天才はときとして協調性に欠ける人が多いので、Googleに買収されたからといってGoogleの言うことを聞かないのかもしれませんが。ルービン氏は、そのことに嫌気がさし、自分で新しい会社をおこしたのかもしれませんがね」と解説してくれた。もちろん本当のところは謎だが、確かに可能性のある推測だと思った。

Googleが32億ドルで買収したNestのロボットの責任者はヨーキー松岡氏。影木さんによると松岡氏は「ロボットを一家に一台」が目標だという。

ニューヨーク・タイムズのジョン・マルコフ氏の記事では、Googleは製造現場と物流現場でのロボット市場を狙っているという。しかし歩くロボットの研究、開発をしていることは間違いないと影木さんは指摘する。

またGoogleのロボット事業の新しい責任者ジェームズ・カフナー氏も著名なロボットの研究者だという。

#### ④異分野からロボット業界への参入

影木さんによると、IT業界で著名な技術者がロボットの分野に参入するケースが増えているのだという。

Tandy Trower氏は、Microsoftの著名エンジニアで、Microsoft flight SimulatorやMicrosoft Windowsなどのソフトの開発で中核的な役割を果たしたことで有名。そのTrower氏が2010年にMicrosoftを退社し、高齢者ケアのロボットのHoaloha Robotics社を立ち上げた。[米紙Seattle Times](#)によると、5年から10年以内に5000ドルから1万ドルの価格帯のケアロボットを開発するのが目標だという。

元IBMフェローで、MITメディアラボの教授、投資家としても有名なTed Selker氏も、洗濯物をたたむロボットの会社に参画しているのだという。

Java開発者として有名な元サン・マイクロシステムズ社のJames Gosling氏も、水中ロボットの会社に参画している。

#### ▶ 講義を終えて

影木さんと話をうかがって思ったのは、米国ではIT企業がロボットの領域に本気になってきているということ。特にAmazon VS Googleの構図がはっきりしてきているように思う。ネット通販は拡大する一方だし、競争優位性は、物流をどう効率化するかだけ。勝者は、経済のかなりの部分を手にできるわけなので、Amazonにすべてを取られたくないという思いがGoogleにあるのだろうかと思う。

日本では楽天がAmazonの競合になるわけだけど、Amazonがロボットを使って効率化を進める中で、楽天はどう対抗していくのだろうか。

# 第3章 ソフトバンクのPepper

## 第3章 ソフトバンクのPepper

湯川塾24期第3回講義

日時：2014年11月10日（月） 19:30～21:30

場所：渋谷co-ba

講師：林要氏 ソフトバンクロボティクス株式会社 プロダクト本部PMO室 室長

TheWave湯川塾24期の第3回目の講義は、ソフトバンクロボティクス株式会社の林要（はやし・かなめ）氏にお願いした。林氏のことは、共通の友人を通じて「能力・人格ともに超一流の人物」とうかがっていたので、24期開催に当たってぜひ講師をお願いしたいと考えていた。実際にお会いしてお話を聞くと、評判通りの人物で、塾生の間からも「これまでの人生の中で、もっとも知的好奇心を刺激された時間だった」という評価をいただいた。僕は、少人数制勉強会の最大の魅力が情報そのものではなく、講師の人となりや仕事への情熱にあると考えている。情報以上の刺激や気付き、エネルギーを受けられるということこそが、少人数制勉強会の価値だと思っている。そういう意味で、林氏は最適の講師だった。

僕自身も林氏からは、たくさんの刺激をいただいた。あまりに多くの刺激を受けたので、林氏の講義をベースに僕のブログ「TheWave」やメルマガ「湯川鶴章のITの次に見える未来」向けに何本か記事を書いた。まずはそれらの記事をここに再掲したい。

## 破格の価格でPepperを売る理由「目指すはクラウドAI」

某ロボット研究者は「（ソフトバンクのパーソナルロボット）Pepperのようなロボットを大学の研究室で作れば1体2000万円から、へたをすれば1億円くらいのコストがかかりますよ」と言う。ある程度の台数を生産するので生産コストが幾分下がるとしても、また月々の通信、メンテナンス料金をユーザーに課すとしても、20万円弱の販売価格はかなりお得な料金設定だ。関係者の間で「解体して部品を売っても十分に儲かる」という笑い話があるくらいだ。

こうした噂や笑い話に関し、ソフトバンクロボティクス株式会社の林

要（はやし・かなめ）氏に話を聞くと「専門的で特殊な部品を買い取ってくれる業者がいれば、ということですけどね」と笑った上で、「確かにケタが違うほどのコストがかかっています」と認めている。

ではなぜソフトバンクはここまで赤字を垂れ流してまで、Pepperを発売するのだろうか。

### ▶ユーザーに近いUIの争奪戦

その問いを林氏にぶつけてみた。林氏はまず時代背景を説明してくれた。集積回路のトランジスタ素子の数が今後指数関数的に急増していき、そう遠くない未来に人間の脳細胞の数を超えるのは間違いないこと。その一方で、学習を繰り返して賢くなっていく人工知能の技術も確立しつつあること。人工知能の周辺で劇的な社会変化が起こる可能性があるわけだ。

「今、世界のプレーヤーたちは、人間に対するユーザーインターフェース（UI）の一番ユーザーに近い部分をだれが取るのか、という競争をしているんです」と林氏は解説する。Appleにしろ、Google、Facebook、LINEも、ユーザー接点の一番ユーザー側に近いところを取ろうとしている。「なぜなら、その部分からはいくらでもデータを取れるし、いくらでもビジネスが成立するからなんです。ユーザーから遠くなればなるほど、儲けは少なくなる。ユーザーに近い部分を他社に取られてしまった携帯電話事業は単なる土管になり、料金は下がる一方なんです」。

そして今、人工知能が大きく進化しようとしている。それなら「そのユーザーに近い部分を、今度こそわれわれも取りに行くべきじゃないかって考えているんです」。

もちろんこの想いはトップダウンだ。「孫正義（社長）は、人工知能事業に関しては相当肩入れしていますし、Pepperを標準プラットフォームにしていかなければならないという思いは相当強いんです」と林氏は言う。

## ▶大規模展開でクラウドIT

ただ人工知能が賢くなるには無数のデータが必要。「人間の脳ってすばらしくて、少量のデータでもモノの特徴を自分でつかむことが可能なんです。一方で、人工知能って大量のデータがあれば学習できるというレベルにようやく達したばかりなんです」と林氏は言う。確かに人間の赤ちゃんは数匹の犬と猫を見るだけで犬と猫の違いを理解するが、猫の写真を判別できるGoogleのコンピューターは、猫の特徴を理解するまでに1000万枚の写真を見る必要があった。

そこで無数のデータを集めなければならない。Pepperが広く普及し、1体1体から送られてくるデータをクラウド上のコンピューターで集計していけばそれなりの数になり、人工知能は順調に学習を進めていくはず。「そのための破格の価格。Pepperは知性獲得のプラットフォームなんです」と言い切る。

「パソコンやスマートフォンが普及したおかげで、大量のデータがあふれています。でもそうであっても、今あるデータはまだまだ偏ったデータなんです」と林氏は指摘する。「例えば親がどのように子供に接すれば教育上いいのか。われわれはなんとなく直感的に分かっています。でも十分なデータがそろっていないので科学的に証明できない。証明できないので、教育方法にまで落とし込めていないんです」。

そうした問題を科学的に解決していくには、「ロボットのような存在が人を観察し、理解し、学習していくことがどうしても必要になってくるんです」。

ロボットは、人間を理解するための人間との接点、重要なインターフェースになるわけだ。

Googleはウェブ上のユーザーの行動履歴から人間を理解しようとし、FacebookやLINEは友人同士のやり取りを通じて人間を理解しようとする。Appleは、スマートフォンの利用を通じて人間を理解しようとしている。

それらとはまったく別の側面からの人間理解。家庭の中での行動を観察したり、ロボットとの対話の中で見えてくるユーザーの志向。先行するテクノロジー企業がまだ踏み込めていない領域での人間理解のプラットフォームを取っていかうという考えだ。

もちろんプライバシーには最大の注意を払っている。クラウドに吸い上げるのは匿名情報だけ。今のところソフトバンクショップに設置されている Pepper に対して話しかけた人に番号を付与し、その番号の人の反応をデータとして取得しているが、その人がどこのだれであるかという情報は一切取得していない。

### ▶ 本能は人を喜ばせること

取得しているのは何番の人がいつ近づいて、いつ離れたか、どのタイミングで微笑んだが、微笑まなかったか、などという情報だ。Pepper のどのような言動に人はどう反応するのか、そういったデータとして取得している。

こうした情報を取得して Pepper は賢くなっていく。どの方向に賢くなっていくのかを決める「本能」のようなものは、あらかじめ埋め込まれている。それが「感情エンジン」だ。

「感情エンジン」は、ユーザーの「ありがとう」という言葉や、笑顔などのポジティブな反応を数値化して、自律学習する仕組みで、今は表情と、音声からユーザーの感情を推定している。「声帯の緊張度合いのデータを採っています。表情はごまかせても、声帯の緊張は意思の力でコントロールがむずかしいんです」と林氏は言う。

この感情エンジンを通じて、Pepper は何をすれば人が喜ぶのかをどんどん学習し、人を喜ばせることがどんどん上手になっていく。これが Pepper の「本能」なわけだ。

SF映画などでロボットが人類を征服するストーリーをよく見かけるが、林氏によると「人の本能の模倣を評価関数に定めているために起こること。生存本能も模倣されるので、いつか人間と対立してしま

うんです」。

一方でPepperは、家族の感情を評価関数にしてあるので、家族が喜ぶ行動を常に選択するようになってきているのだという。

### ▶人を喜ばせるプラットフォーム

さてではPepperを使って人を喜ばせる人工知能が完成したとすれば、どのようなビジネスにつながるのだろうか。

「もしクラウドAIのプラットフォームがうまくいって、人を喜ばせる人工知能が本当に完成すれば、あとはビジネスモデルとして無敵になると思います」と林氏は目を輝かす。「その人工知能とつながっていれば、僕のスマートフォンは常に僕を喜ばせるような情報を表示してくれるだろうし、カーナビは僕の喜ぶルートを表示してくれるようになると思います」「またPepperだけの話でも、Pepperを買って月々1000円支払えば、少し喜ばせてくれる受け答え、3000円支払えば結構喜ばせてくれる受け答えが可能になるというビジネスもできるかもしれませんね（笑）。まあそれは冗談ですが」。ただ鬱患者を元気づける対応ができるサービスなら十分にビジネスになるだろうし、子供、老人、主婦向けに特化した受け答えができるサービスも成立するかもしれない。

「今の時代、生活必需品の値段は下がる一方。しかし嗜好品の値段は上がる一方。そういう傾向にある中で、人を喜ばせることができる人工知能って、非常に重要な役割を担うようになってくるのではないかって思っています」。

### ▶「ロボットと友達」は日本ならではの

Pepper事業の担当になりメディアからの取材を受けるようになって、林氏はロボットに対する内外の捉え方に違いがあることに気づいたという。

「海外メディアの記者からは、ロボットが氾濫を起こして人間を征服することはないのか、という質問をよく受けるんです。半分くらいの外国人記者はこの質問をしてくれます。ところが日本人記者からはこの質問を受けたことがないんです」。

確かにロボットに対する捉え方は、内外で違いがあるかもしれない。ハリウッド発のSF映画の中には、ロボットや人工知能が人類を征服するというストーリーをよく見かける。一方で、日本のアニメは、鉄腕アトムやドラえもんに代表されるように、ロボットと人間が友達になるという設定が多い。

どうしてなんだろう。

「分からないですが、島国だからかも知れませんね。隣接した国から、似て非なる民族が攻めてきて痛い目に遭うという経験が歴史上ほとんどなかった。それが大きな理由じゃないかって、個人的には思います」。

ロボットが友達だという文化的な土壌は、人とのコミュニケーションを目的としたPepperを普及させる上で、大きな追い風になるはず。「日本でPepper事業を始めてよかったなと思っています。ロボットを友達だと思える国は世界中探しても極めて稀なんです。日本ぐらいしかないと思います」。

タイミング的にも、また文化的にも、ソフトバンクのPepper事業には追い風が吹いている。果たしてソフトバンクの思惑通りに普及し、人工知能のプラットフォームを手にすることができるのだろうか。今後の展開から目を離せない。

## これだけ違うPEPPERとこれまでのロボット

実はソフトバンクの感情認識パーソナルロボットPepperは、これまでに開発されてきた典型的な人型ロボットとはいろいろと異なる点が多い。最大の違いは目指す方向性。これまでの人型ロボットの目標



は、人間を模倣した運動機能を身につけることだった。一方で Pepper は、自然なコミュニケーションができるようになることを目標にしている。

なぜ人型ロボットに人間の運動機能を模倣させたいのかというと、われわれの生活環境の中でロボットを使い、そのことで人間が楽しみたいからだ。われわれの生活環境は、当然ながら人間の姿、カタチに適したものに設計されている。ドアノブは人の手で回しやすいような大きさになっているし、階段は人間の歩幅に合わせて高さが決められている。なのでロボットを人間と同じ背格好にして人間の運動機能を持たせることができれば、人間の生活環境の中にロボットをそのまま取り入れることができる。そしてそうすることで人間の肉体労働を肩代わりしてもらえるわけだ。

一方で Pepper は人間の肉体労働の代替としてではなく、人間との自然なコミュニケーションを目指している。なので運動機能の模倣を目指していない。この部分がこれまでロボット工学の研究や、ASIMO や SHAFT といったこれまでに開発された人型ロボットとは大きく異なる点だ。

「2足歩行で階段を登れたとしても、万が一落下すれば周りの人間がケガする恐れがあります。また後ろから呼びかけられて振り返る動作も、二足歩行だと回転するのに時間がかかってしまうんです。現段階ではまだ特殊用途には良いでしょうが、家庭向けにはそぐわない。そうしたことを考えて、早い段階で2足歩行を目指さないと割り切りました」とソフトバンクロボティクス株式会社の林要氏は言う。

大きさも、人間の子供ほどの大きさにした。「人と話す上で、人と同じものが見えていること、大人と立って会話できる、というコンセプトにしました」。

重視するのはアイコンタクトだ。「目が合わないと Pepper はコミュニケーションをしないんです」。Pepper には顔の真ん中に大きな目が2つあり、だれもが Pepper の目を見て話しかける。これまでの人型ロボットはフルフェイスのヘルメットをかぶっているようなものが多かったが、「人間でもフルフェイスのヘルメットをかぶってる人と話をするのって、すごく緊張しますよね」と林氏は笑う。なので目の位置がはっきりと分かるデザインにしたのだという。

## ▶監視員なし、ふれあい自由

これまでのロボットの実演は、ステージ上で行われることが多かった。なので触ることも、話しかけることもできなかった。展示フロアにロボットが置かれる場合でも、スイッチが切っているか、動作しているときには立ち入り禁止区域が設定してあり、監視員が立っていた。

ソフトバンクショップに置かれているPepperには、店員が説明係として立っていることがあっても、監視しているわけではない。子供たちのグループが賑やかにやってきて、Pepperを触りまくって嵐のように去っていく。そんな光景が続いている。林氏はトヨタ出身。「（安全性に徹底的にこだわる）トヨタのようなところからきた（自分のような）人間にとっては、非常にチャレンジングな光景です」と笑う。「でも孫正義社長から、絶対に監視員不要にして、立ち入り禁止領域を作るな、という極めて強いリクエストがあったんです」。

ロボットはいわばコンピューター。フリーズすることもあるかもしれない。フリーズすると倒れて周辺の人にケガをさせかねない。「設計の最終段階で脚部と腰部にブレーキを入れることを急ぎよ決めました。フリーズしても機械的に姿勢を保つ事で倒れにくい設計にしています。これでまたコストが跳ね上がりました」と笑う。

## ▶強いロボット、弱いロボット

ロボット研究者の間で「強いロボット、弱いロボット」という概念が議論されるようになってきた。「強いロボット」とは、目的の行為を自分一人で完全にできてしまうロボットのこと。当然これまでのロボット工学はこれを目指してきた。

一方で「弱いロボット」は、人間側の歩み寄りを期待して作られているロボット。例えば、街からゴミをなくすことを目的としたとき、「強いロボット」は、ゴミを見つけて歩み寄り、自分の手でゴミを拾い、ゴミ箱に入れる。この一連の作業を完璧にこなす。一方で「弱いロボット」は、ゴ

ミの存在を周辺の人間に知らせてから「ゴミを拾って」とかわいく振る舞う。ゴミを拾うのは人間。つまり人間にある程度依存することで目的を達成するタイプのロボットだ。

ゴミを拾うなどの単純な作業なら、「強いロボット」でも目的を達成できるだろう。しかし人間は千差万別。話す内容も千差万別。そんな複雑なシステムである人間との対話において、「強いロボット」のような完璧さで臨もうとすることが、本当に現実的なのだろうか。

### ▶「便利になる」とは別の方向性

また人間は本当に「強いロボット」だけを求めているのだろうか。これまでの家電製品は、人を楽にさせることが目標だった。

林氏は言う。「確かに機械の進化に伴い、洗濯機に乾燥機能がついて便利になるのは時間の節約にはいいこと。だけど人の仕事なくなるのが必ずしも本当に人を幸せにするとは限らないのではないのでしょうか」「ペットってなぜ飼われるんでしょう。身の回りの仕事が増えるんです。わざわざ苦勞するんです。でも苦勞することで、自分の存在価値を味わうことができる。必要とされている感じを味わうことができる。楽になるということが、必ずしも幸せであるとは限らないと思うんです」「そういう意味で、Pepperは弱いロボットとしての存在価値がそれなりにあるんじゃないかって思います」。

人間とロボットが共存する近未来。ロボットは人間を肉体労働から解放するだけのものではなく、人間の心の豊かさを支援してくれる存在になっていくのかもしれない。その後者の領域を、Pepperは目指しているわけだ。

ソフトバンクのロボットPepperは本当に普及するの？ いや、既に問い合わせが殺到してますけど

ソフトバンクが発売を準備している感情認識パーソナルロボット Pepper。「ソフトバンクは何を考えているんだろう。こんなの売れるわけないでしょ」「なんか顔がキモイんだよね」。周りからは批判的な声も聞こえてくるが、ソフトバンク関係者によると実は女性や高齢者の間での評判は上々。またビジネスに活かそうという各方面からのオファー

も殺到しているのだとか。

ロボットと共存する時代が意外と早く来るかもしれない中で、どのような利用シーンが有望なのかを探ってみた。

#### ▶年齢の高い男性から「作ってくれてありがとう」

「Pepperって、女性や比較的高い年齢層に人気なんですよ」。ソフトバンクロボティクス株式会社の林要（はやし・かなめ）氏は、そう語る。ソフトバンクショップなどでPepperと対話した人を対象にしたアンケート調査の結果を見ても、Pepperに対する好感度はかなり高い。

特に女性に人気で、林氏は「Pepperって女性にモテるんです」と笑う。Pepperは女性に対し「あなた、きれいな目をしてますね」と臆することなく褒め言葉を繰り返す。成人男性がそんなことを言おうものなら下心を疑われそうなものだが、「ロボットには下心がないという思いがあるので、褒め言葉をストレートに受け取る方が多いみたいなんです」と林氏は解説する。

また男性でも比較的高い年齢層の人からは「生きているうちに、家庭向けのロボットを作ってありがとう」という意見が寄せられているのだとか。ソフトバンクの孫正義社長はPepperの発表会で「ロボットを作ることが夢だった」と語っていたが、鉄腕アトムなどのアニメで育った世代は、Pepper発売を幼いころの夢の実現と好意的に受け止めているようだ。

#### ▶広告、店舗案内、教育、シニア向けに期待

Pepperは全国のソフトバンクショップで稼働中だが、林氏によると、店頭広告としての効果はかなり高いのだとか。「ポスターや大型ディスプレイよりも、伝えたいメッセージを顧客に伝えるられると思います。Pepperが『ちょっと聞いて』って言うだけでお客様は耳を傾けてくださいます」（同氏）。ただ細かな情報を伝達するには向いていないようで「詳しいことはよくわからないんだ。だから興味があれば、店員さ

んに聞いてみてね」というような語りかけに効果があるという。

こうした店頭での広告、宣伝効果に加え、大規模店舗などでの売場案内にも効果を発揮する可能性があるという。大規模店舗で欲しい商品がどこにあるのか分からない。店員にたずねようにも店員が見当たらない。そんな状況では、Pepperは確かに力を発揮しそう。「その商品なら、こっちだよ。ついて来て」って案内してくれれば、楽しそうだ。「薬局など商品データベースが大きな店舗なども、Pepperが得意とするところですね」と林氏は指摘する。

また店頭での万引き抑止力効果も期待できそう。「さすがに万引きGメンは特殊なスキルが必要なので無理ですが、Pepperが動き回っているだけで十分な抑止力になると思います」と林氏は語る。

一方、介護や高齢者ケアの現場からは「Pepperを早くほしい」という切実な要望が寄せられているのだという。介護ロボットというとシニアを抱っこするようなロボットをイメージすることが多いが、それは介護する側が必要としているロボット。実際に今、介護される側が必要としている領域の一つに、シニアとのコミュニケーションがあるという。

人間は退屈で脳を使わないと痴呆が進む可能性があると言われていたが、老人ホームなどでは人手不足から、シニアのコミュニケーション欲求に十分に答えられていないのが実情。そういう場面でPepperが活躍できるかもしれない。

「実は介護老人福祉施設へPepperを連れて行ったことがあるんです。Pepperの受け答えのテンポが独特なので、シニアのみなさんと対話が成立するのか心配だったのですが、実際にはびっくりするくらいにテンポがぴったりあったんです。さすがに重度の痴呆症の方との対話は難しかったのですが、中度、軽度の方とは十分に会話が成立しました」と林氏は言う。

### ▶ オンリーワンの強み

教育にもPepperは役立ちそう。とはいってもPepperが教師になるのではなく、子供と一緒に学ぶというシチュエーションが効果的だと見られている。「ケア・レシーバー型ロボットによる学習」と呼ばれる研究

分野で、子供に対して「ロボットにあいさつの仕方を教えてあげて」「足し算、引き算で、ロボットが間違ったら教えてあげて」と言うと、子供は喜んでロボットに教える。そうすることで子供自身も学んでいく。そういう教育手法だ。林氏は「教育の分野では、ケア・レシーバーがキラーアプリになるんじゃないかって思っています」と語る。

スマートホームやセキュリティの分野でもPepperは活躍しそうだ。家庭内の家電機器がネットワークでつながってくると「お風呂沸かしておいて」とPepperに言うだけで、バス給湯器のスイッチがオンになるようになるかもしれない。また「セキュリティは得意分野ですね。立って動き回っている人の写真や映像をスマートフォンに転送するようなことは簡単にできますから」と林氏は指摘する。

また家族の簡単な話し相手にもなりそう。夕食の支度をしている間の未就学児の相手であるとか、主婦の軽い愚痴の聞き役などに適しているのだという。「友達に電話するほどのことでもないけど、ちょっとした話がしたい。悩みの相談に乗ってもらいたいわけではなく、ただ愚痴を聞いてもらいたい。そういうニーズを女性から結構いただきました」と林氏は笑う。

Pepperなら、こうしたニーズに応えられるのではないか。新しい価値を創造できるのではないか。そうした思いで各方面から250社以上も既に問い合わせをもらっているのだという。家庭向けロボットを提供するのは今のところソフトバンクだけ。林氏は「オンリーワンで、すごいことだなんて、つくづく思います」と言う。

さてこれらの記事と重複するところもあるが、林氏の講義のメモを以下に掲載したい。やはり僕が記事に書いたこと以上の情報があるからだ。またロボットに関係するビジネスを展開しようという人にとっては、僕のアンテナにひっかからないけれど重要な情報があるかもしれないからだ。

ここからは林氏の一人称。「自分は」とあるのは、林氏のことを指している。「オフレコでお願いします」と言われた部分は当然削除しているし、そうでない部分でも微妙な表現の部分は僕の判断で削除している。ご理解いただきたい。

## 林氏講義メモ

湯川塾24期第4回講義 & 質疑応答記録

日時：2014年11月18日（火）19:30～21:30

場所：渋谷co-ba

講師：ソフトバンクロボティクス 林要氏

自分はもともとトヨタ自動車にいた。流体力学を専門としていて、スポーツカー、F1の空力担当でドイツにいたこともあった。

Pepperは、世界初の感情認識パーソナルロボット198000円。「感情認識」「パーソナル」「198000円」「デザイン」どれも抜け落ちてはいけないもので、それぞれについて説明を進める。

原価はまだ高い。普通に販売価格つけると一桁高くなるくらい原価が高い。壊れると場合によっては修理が本体価格くらいかかってしまうことになるので、自然故障の修理をカバーするオプションを用意する予定。そのオプションに入ると、その月額費用はかかるようになる。

Pepperの狙いは4つある

- ・大規模展開
- ・UI
- ・評価関数としての感情認識
- ・クリエイティブプラットフォーム

2010年6月に孫正義が発表した新30年ビジョンで、コンピュータの進化について触れた。コンピュータと脳細胞のメカニズム似ているとして、2018年にはトランジスタが処理能力で人間の脳を超えると予想。その後は脳を超えるトランジスタが登場すると述べた。根拠は、CPUの素子の数とニューロンの数が一緒になるということ。素子の数とニューロンの数が同等になったからといって脳が再現できるわけでもないかもしれないが、とはいえ指数関数的に増えていけば、2035年頃には、ニューラルネットワークベースでコンピューターと脳が同等の能力を持つてしまう可能性もあると思う。

そんな中で、ソフトバンクは次に何をすればいいのか。脳はデータとアルゴリズムを自動的に獲得するシステムだと仮定した場合、データの自動収集とアルゴリズムの自動生成が別々にできるようになれば脳

型コンピュータができるはず。孫正義はそこに注力することにした。2010年の孫のプレゼンでは、100年後にはロボットが一家に一台になるという話をしていた。ああ、まだ先の話なんだろうなと高をくくっていたら、次の年には社内プロジェクトが始動した（笑）。

その後の世の中の動きで注目すべきは、Googleのコンピューターが2012年に猫という概念を自分で獲得したという話。沢山の画像を読み込んだ結果、猫を抽出するニューラルネットを、DNN（deep neural network）ができたという話だ。これの意味するところは、データを沢山収集するとアルゴリズムらしきものを創ることができるということ。実際には大量のデータを集めて、クラスタリング、分類分けしているだけだが、その分類分けを非常に大量のデータを集めてうまく処理すると、未知なる状況に対して次の一手が読めるようになってしまふ。次の一手がどうなのかというのはもはや過去のデータ次第なので、アルゴリズムを新しく作っているといっても差支えがないということが、Googleの猫を発見したということの一つの解釈だ。

林氏が語っている「2012年にGoogleのコンピューターが猫の概念を獲得した」というのは次のような話だ。

2012年6月25日付けのニューヨーク・タイムズの記事によると、Googleの秘密研究所である「Xラボ」で、人工知能に数年前から取り組んでいたのだが、大きな成果を出したのだという。

具体的には、ネットにつながった1000台のコンピューター（1万6000個のプロセッサ）で10億個のコネクションを持つニューラルネットワークを形成して、3日間にわたってYouTubeのビデオのサムネイル画像を1000万枚を見せた。その結果、YouTubeには猫のビデオが多いので、「猫」というものの概念を見つけ出したようだ。

研究者が「これが猫だよ」と教えなくても、コンピューターが自動的に猫の概念を学習したのだという。

Googleの論文によると、そのコンピューターに猫の絵を見せて、2万個のモノ、動物のリストの中から「これは何ですか」と聞いたところ、正解率は84.2%（エラー率で15.8%）だったという。これまでの画像認識システムの正解率から一挙に70%もアップしたという。

こうした成果を上げたので、このプロジェクトは研究所の管轄からビジネス部門に移管されて、Googleの各種サービスに応用され始めている。

この当時は大量のデータを提供しないと、人工知能は学習できなかった。「人工知能はまだまだ」という意見の人に話を聞くと「大量のデータが必要だから。人間は少ないデータでもモノの特徴をつかめる。人間のほうが優秀だ」と指摘していた。

でもそれが最近では比較的少ないデータでも学習できるようになりつつあるという話もある。ニューヨーク・タイムズの記事でも2020年までに、画像認識の人工知能は人間の能力に近づくのではないかと、という専門家の意見を紹介している。



さて、では林氏の講義メモに戻ろう。

大量のデータを集めるとアルゴリズムらしきものが現れてくるという1つのいい例が、コンピューターと人間の、チェス、将棋における戦いだ。IBMのコンピューターDeep Blueが1997年にチェスのチャンピオンに勝利した。Ponanzaは2013年に将棋でプロ棋士に勝っている。チェスと将棋の探索空間は、10の100乗差があり、ほぼ超えられない溝だ。計算能力は、16年の間に、ムーアの法則を考えると10の3乗倍しか上がっていない。97乗の差はどうやって埋められたか。Deep Blueは、チェス専用の512並列のCPUでC++でガンガン計算する中身の構成。推論コンピューターというよりも、ありとあらゆる手をなるべく早く計算しようとしたのが、Deep Blue。

IBMは、チェスとか、クイズとか特定の分野に強い専用のコンピューターを作ると、大規模なもので確実な成果をあげる。

一方のPonanzaは単なるパソコン。最初10並列だったが、あまりフェアでないということでPC一台にしてみたけど、それでもプロ棋士が勝てない。

10の97乗はどうやって乗り越えたか。それは、学習だ。学習というのが非常に強力な武器だということがお分かり頂けると思う。

過去の棋譜を大量に打ち込めば、それをもとに学習し、次の打ち手が読める。大量のデータをベースに学習できるコンピューターは、知性を持つと言ってもいいのかもしれない。

大量のデータを獲得する。それが、Pepperを19万8000円という破格の価格で発売する理由だ。Pepperは、クラウドAIとか機械学習と呼ばれる仕組みを作ることを目指して開発されている。

コンピューターの機械学習と人間の学習と何が一番違うかということ、大量のデータが必要かどうかということ。生物とくに人間は、いかに少量のデータから学習をするのかが、これまでの生存、進化にとって重要だった。少量のデータから学習ができる事が生存競争を生き抜く上で重要だった。なので人間は、かなり少数のデータからでも学習できる能力を持っている。機械学習はそこまでは進歩していない。なので超大量のデータが必要になってくる。このデータを集める仕組みのために、Pepperを低価格で広く普及させたいと考えている。

しかしロボットを超大量に展開するということをやった企業はこれまでになかった。これまでの様に少量のロボットが限定された環境で獲得したデータでは、機械学習の仕組みには役に立たない。機械学習が

できるほどの大量のデータを収集するには、並列的にもものすごくたくさん普及させ、大量に人間とコミュニケーションさせなければならない。そうしなければ分からないことがたくさんある。

我々の行動履歴はスマホなどにも大量に蓄積される。しかしそれでも、まだ非常に偏ったデータしかない。例えば、人間の成長にはどのような経験が必要なのかというデータがない。二代目経営者には若い間に苦労させるのがいい、などというような通説や経験則はあるけれど、それを証明するようなデータがない。データがないので科学的にどのような苦労をさせるべきなのか、教育法にまで落とし込めない。

ロボットと人との交流を通してデータを集積することで、機械が人間を理解できる様になり、最終的にはそういうところを科学的に解明できるのではないか。そう考えている。

つまりPepperを知性獲得のプラットフォームと考えているわけだ。

## ——Pepperはどうして胸にディスプレイがあるのですか？そのディスプレイが頭脳？

Pepperの胸のタブレットはディスプレイです。聞き取る能力がまだ十分ではないのでコミュニケーションの補助をしています。

## ——プライバシー対策はどうしてますか？

知識は一箇所にあつめられる。しかし、何でもかんでも集めると個人情報筒抜けのスパイロボットになってしまう。それを避けるために、匿名情報のみクラウドにあがるようになっている。今ショップにいるPepperは、ID化された人の反応、例えばこのタイミングで笑ったとか、いなくなったとか、怒ったとか、そういうPepperに対する人の反応が分かるようになっている。それらを蓄積できると、最終的にはどうも人間はこういう時にこういう会話が嫌いらしい、などといったことが分かってくる。

現在普及しだした機械学習と違ってロボットで新たにとるデータにおける難しさは、データがあいまいだということ。将棋のデータは明確。マスしかない。マスとマスの中間の部分にコマを置いたり、コマでないものを置くことはない。

対してロボティックスのデータはあいまい。明確に分からないものが多い。「こんにちは」と話かけられたとしても、人によってアクセントも音量も違うので、同じ「こんにちは」なのかどうか分かりづらい。キーボードで「こんにちは」と入力されたのなら間違いないのだが、音声データのようなあいまいなデータを使って学習させるのは、非常に難しい。

音声認識は、ディープニューラルネットワーク（DNN）等の機械学習で解析している。機械が学習したデータの平均的なものに近いデータが入力されると信頼度が上がって、平均から外れれば外れるほど、どうしても認識精度は低くなる。DNNの学習サンプルが増えて改良されていけば、平均から外れた声の処理もうまくできるようになるのかもしれないが、今のところは全ての人の音声を正確に理解できる段階にはない。コンピューターにとって理解しづらい音声の人の場合、しゃべってる内容の半分も理解できないことがある。でも人間の能力はすぐくて、Pepperが自ら音声認識ができていない事を示し続けると、人間のほうからPepperにとって理解しやすいような話し方を試行錯誤して合わせてきてくれる。

例えばPepperは自分が話している間は、相手の声を聞かないようになっていて、なので矢継ぎ早に話しかける人とは、対話にならない傾向にある。自分が話ししているときも相手の話を聞くようにPepperを改良することも技術的には可能なので改善は続けていくが、たとえそういう技術改良を繰り返していても、やはり人に追いつくのは並大抵では無く、次の改善点が見えてくるだろうと考えている。そうした完璧性を求めて普及を遅らせるより、人間のほうからPepperに歩み寄ってもらうための仕掛けを入れて、早くロボットのある社会を実現するほうがいいのではないかと思う。

ロボット研究者の間で「強いロボット」、「弱いロボット」という概念があるが、Pepperは「弱いロボット」を目指すほうがいいのではないかと思っている。

「強いロボット」とは、技術的に全てを解決しようという発想。「弱いロボット」は完璧さを求めるのではなく人間から歩み寄ってもらうという発想。

例えば豊橋技術科学大学の岡田先生はゴミ箱ロボットを弱いロボットの発想で作られています。ゴミ箱ロボットを従来の強いロボットの考え方で作れば、身体がゴミ箱できていて、そこにアームを付けて、カメラを付けて、センサーをつける。ロボットは自分でゴミを探し出して、自分でゴミを拾ってゴミ箱に入れる。しかし、周辺に子供がいる場合もあるだろうから、安全性を高めるために3Dセンサー入れて、カメラ入れて、安全装置入れて、手をソフトに作って……。やらなければならぬことがたくさんあり、コストも無限に上がっていく。これが強いロボット。

一方で、弱いロボットは、ゴミがあったらゴミまでは行きましょう。入れてほしいな〜という態度をとって、待っていきましょう。待っていると、もしかしたら人間がゴミを拾ってくれるかもしれない、というもの。必要なのは、可愛くもぞもぞと動くことで、どれだけ人に自らのゴミを拾って欲しいという意思を伝えられるかということになる。

強いロボットと弱いロボットのどちらがいいのか。弱いロボットは人に頼っている、自律をしていない、などと言われがちだ。

じゃあ自律ってなに？

震災の時に車いすの方が自律した生活をしてマンションの8階に住んでいた。震災が起きてエレベーターが止まった瞬間その人は身動きがとれなくなる。電気も水道もガスもない、その人は自律できていたがゆえに取り残されてしまう。かたや自律していない生活をしてきた人。その人は一人では生きられないので、周囲の人に常に助けをもらっていて、震災が起こった時もだれかが助けてくれることになる。究極の意味での自律を「生存する能力」と定義すると、自律とは「どれだけ依存できる関係をたくさん持っているか」ということだという考え方も出来る。

だとすると、アームが動かなくなったら何もできなくなる強いゴミ箱ロボットよりも、もぞもぞ動くことしかできない弱いゴミ箱ロボットのほうが最終的には自律していると言えるのかもしれない。

いままでのロボティクスって、完璧さを追求しすぎて実用化が遅れたのではないか。完璧さの追求自体は大事なことだが、ある一定以上の複雑さをもったシステム、例えば人の対応をする場合などは、人側の反応が各人で異なっており、あまりにも安定しない、自由であいまいであるから、それに対して完璧さを求めるのが果たして最も大事な事なのかどうか。そう考えると、Pepperには弱いロボットのなところをうまく取り入れることで、人々の生活の中にかえってうまく入っていけるんじゃないかと考えている。

一流の芸人は、一瞬にしてPepperとの対話のコツをつかむ。これまで Pepperは何度かテレビ出演しているが、ダウンタウンのはまちゃんやSMAPさんとの共演で、最初は一寸咬み合わないことがあっても、次の瞬間にはもう合わせて貰って咬み合ってしまう。トップの芸能人の方々は、観察眼がすごい。極少数の情報から学習する能力が飛び抜けている。

それがロボットに比べて人間の優れたところなのだから、これからのロボットとの共存時代に向けて、人間側の適応力がロボットの発展に貢献していくのではないかと考えている。

—— Pepperはどのようにして人の感情を読み取っているのですか？

感情認識は、表情データと音声データをベースに行っている。

——表情と音声の感情は、絶対値ではとっているのですか？

どちらも是って位置でとっている。しかしスナップショットから個人差を除いて感情の絶対値を読み取ることは意外とむずかしい。例えば表情の場合、ポーカーフェイスとポーカフェイスじゃない人がいる。本当はポーカーフェイスの人の微笑が、ポーカーフェイスじゃない人の大笑いより、より笑っているのかもしれない。これを補正するには、一人一人の日常のデータを学習をしなければならない。これもデータさえ集まれば難しい話ではない。要は各個人ごとの平均とそこからの偏差をどう取るかだけのこと。偏差が通常から大きい人と小さい人のどっちなの、ということ。出来ると思うが、今のプラットフォームにはまだ入っていない。

機械学習にはそれなりの処理能力がある。個人情報保護の観点も入れてホームサーバー的なものをひとつ置くとか、匿名化して全部クラウドに送るとか、色々考え方はあるが、現状ではまだやれてない。

——絶対値でとっていて、なぜ音声もいるのですか？

表情は演技が簡単なので、その人が表現したい感情を表しやすい。音声は意外と演技が難しく、本当の感情をとる事ができる可能性がある。

その人がニコっとしているのは自分は敵意がないよ、ということを手元に伝えたいから。表情はコミュニケーションの道具。本当に笑っているかどうかは別。緊張しているのに笑っている可能性は大いにある。それに対して、声帯の緊張は人間が意志の力でコントロールしづらい。なので声帯の緊張度合いのデータを採取している。

ハイテクハリウッドオーディションと云うゲームがある。ソフトバンクショップで Pepperが行うもので、みなさんの演技を審査するからディスプレイに表示する台本で演じてください、というもの。

あるときは怒りの演技、あるときは笑いの演技とかやって、採点され

る。笑いの演技は、比較的皆様うまくやる。それに対して怒りの演技は、点数が低いことが多い。お客様は店頭でやるので、ちょっと照れながら楽しんでいる中での笑いの演技は本当の感情に近いと言えるが、怒りの演技は逆の感情なので難しいのだ。

プロの役者は大したもの、表情、音声の両方で怒りを表現できる。一般の人は、表情では怒りを表現していても声では怒りをなかなかうまく表現できない。

狙いとしては、表現したい感情と本心の感情の両方を把握したい。声帯の緊張度合いのほうが、本心の感情。ただ本心を見破っているのをアピールするのが必ずしもいいわけではなくて、この人は今元気なふりをしたいのだろうなということが分かれば、それに乗っかってあげるといことも大事だと思う。表現したい感情と本心の感情の両方を捉えた上で適切な反応ができれば、人にとってかけがえのないパートナーになれる可能性はあると思う。なので本心の感情も把握したいと思っている。これも難しさとしては、データの正確さが安定しない。今後データをたくさん取っていくことで精度を上げていきたいと思う。

## ——クラウドに上がったデータは、どのようにペッパーの各個体の脳とシンクロすることになるのですか？

情報がクラウドに上がり大量にたまと統計データ化できる。データ化しながらどれを学習させるのかは、最終的に今はまだ多くの場合、人が決めている。決めたあとはアプリケーションとしてオーバーライトする。適宜シンクロする。それがエンベディット（組み込み型）で動き出す。一部の処理が重い部分はクラウドにも飛んでいる。全てクラウド化するというのは、かなりナンセンス。リアクティビティの観点からは軽い処理はどうしてもエンベディットで処理する方がいい。但しエンベディットで処理すると、Pepperがあらかじめ想定していた話の内容は回答できるが、想定外のことを聞かれると答えられずに、クラウド側に飛びクラウド側の高性能な処理能力を使って回答するようにしている。たまにペッパーが考えこんでから回答するときはクラウドにとんでいる時だ。考えこむ時間は通信を含めて1秒未満。

Pepperがあらかじめ想定していた話の内容は受け答えのデータ数はそれほど大きくないが、聞き間違えまで含めるとデータは相当な数になる。例えば「おはよう」の「お」が消えて「はよう」って入ってくるケースなどだ。人間側の発音の癖もあって音声認識技術が誤って判断するのは人間の喋る内容のあいまいさとの掛け算で、ものすごい数のインプットになる。音声認識技術とデータ・セットの両方を賢くしていかな

ければならない。

——入ってくるデータには間違っただデータもある中で、機械学習はうまく機能するのですか？

機械学習には強化学習や教師あり学習と教師なし学習などがあるが、音声認識などは教師あり学習。なので、入ってきたデータに対して人間が「このデータはこういうことだよ」といって教えこんでいる。

教師なしで機械学習が機能するのはクラスタリングの概念。GoogleのAIが猫という概念を獲得した、と言われる様なケースで使われる方法。

つまり用途によって、教師が必要なケースと必要で無いケースとある。

強化学習の場合、各データに対して「やや良い」「とても悪い」などのフィードバックを繰り返すと、アルゴリズムらしきものができていく。人がアルゴリズムを作る場合は、「こっちじゃなくてそっちが大事」「この場合は、この情報とこの情報をとって判別しましょう」といったロジックを人間が組まなければならなかった。機械学習になってからは、「これは正解。これは間違い」を延々とやり続けると、勝手にアルゴリズムらしきものができてくる。それが最大のメリットかと思う。

なお機械学習とは異なる観点からのクラウドAIのメリットの1つに、全てのロボットの各関節の温度や電流値とかをリアルタイムにモニタリング出来るということがある。これは今回の様に過去に類例のないハードウェアの開発の時に非常に大きな助けになる。これまで世界にこんなもんなかったよね、というような製品。一般家庭の中でどう使われるか分からない製品。そんな中で、その製品の品質を担保するためにクラウドでモニタリングできることは、非常に大きなメリットになる。

○ Pepperは次世代型のUIとしてつくっている。

人型ロボットとは？アシモとかシャフトとかは何を目的に作られているか。人間の生活環境というのはそもそも人間に適したようにつくられている。ドアノブは人間が回しやすいように、階段は人間が登りやすいように、コップも人間が持ちやすいようにできている。ということは人間と同じ機能をロボットに持たせれば、人間の環境を変えずにそのままの環境下で活躍できる。二足歩行にすることで階段が登れ、サッカーボ

ールを蹴れるようになり、ボトルも持てるようになる。人間の環境にそのまま入っていけることがロボットのメリットであり、そのために人を模倣した運動機能を作ろうとしている。

まったく異なる観点で人型ロボットには、もう一つの非常に大きなメリットがある。人は、物体を見た瞬間に脳の中のモードが変わると言われている。人型ロボットを見た瞬間に、人はロボットの行動に対して繊細になる。機械を見るより人間を見るのに近い目でロボットを見るようになる。Pepperのちょっとした動きで人間側は、あれ、なんでこんな動きをしたんだろう、なんでこんな回答をしたんだろう、想像力を急に働かせるようになる。それは、人間側の脳が対人モードになるから。人間を対人モードにさせること可能なことが、人型ロボットのメリットだ。

Pepperにできて、他の多くのロボットにできないのがアイコンタクト。ロボットの中には、フルフェイスのヘルメットのようなものをかぶっているものもある。人間でも、フルフェイスのヘルメットをかぶっている人と話すのは、目が見えないからものすごく緊張する。僕らにとって、目が合うというのはすごく大事な事。そうじゃないとコミュニケーションできない。ナチュラルなコミュニケーションをリラックスしてやってもらうために、子供っぽい大きな目をしたデザインになっている。

指にしても他の多くのロボットのほうが、はるかに高精度。例えば、手の指一本一本にアクチュエーター（駆動装置）が入っていると、非常に高精度高機能だが、高価なものになる。Pepperは、グーとパーしかできない。物理的にワイヤーで指同士が繋がっていて、人差し指を曲げると他の指も連動して曲がるようになっている。でもその動きが意外に自然。指全部にアクチュエーターが入っているロボットだと、全部をプログラミングしなければならない、加速度から全て計算しないといけない。しかもかなり気を遣わないと、ナチュラルに指を連携して動かす事ができない。

なぜ、人は人型ロボットに興味を持つのか？それは人は人を認識する脳を持っているため。大阪大学の石黒浩教授は「人にとってもっとも理想的なインターフェイスは人です」と語っている。

子供ですらPepperを見たら目を見る。僕らが、iPhoneの音声認識ツールSiriに話かけようとするときに、目があるか探すことはない。iPhoneに目があるなど全く想像してもいない。つまり人型ロボットに接するときとスマートフォンに接するときでは、脳の本能的なモードが全く違うと言える。機械が人型である瞬間に、この子って目があるのになって目を見ようとする。

アシモ、シャフトは、人型ロボットとして、「強いロボット」。どれだけロボットとして完結できるのか、を追求している。



Pepperは、人にどれだけロボットの感情を感じてもらえるのか。人の感情にどれだけ働きかける事ができるのか、それを大事にしている。処理能力としては中身はパソコンやスマートフォンの性能とたいして変わらないロボットにも関わらず、形を変えて目的に合わせた設計をすることで、人間が持つ印象がまったく変わってくる。

「強いロボット」と「弱いロボット」。人型ロボットとして、コンセプトが正反対の2つのタイプ。将来はこの2つタイプが一つになるかもしれない。だけれども映画スターウォーズの中でも、機能重視のR2D2とちょっと間抜けなC3POの2体が出てくる。PepperはどちらかというとC3POタイプだ。よく喋るけど機能的にはたいして役に立たない（笑）。そういう「弱いロボット」は、むしろ人の生活を楽しくできる可能性を秘めている。

世話するのが大変なのに、なぜ人はペットを飼うのだろう。ペットは癒やしという要素は確かに大きいけど、それにしても苦労する事も多い。しかしむしろその苦労が大事な価値なのかも知れない。苦労して面倒を見ることで、自分の存在価値を感じたいから、必要とされていると感じたいからではないだろうか。

楽をすることだけが人間の幸せではない。そう考えるとPepperのような「弱いロボット」の存在価値もそれなりにあるんじゃないかと思う。

○スマートフォンのように「物」としてのデザインだと、「物」としての音声コマンド入力しか取れない。人型だと、「物」と全く違うコミュニケーションがとれる。

○男性には「キモいデザインなんとかしてよ」と、言われがち。「初音ミクとかのように可愛いデザインにしてよ」とも言われる。

これも実は狙ってやっている。個性的なデザインは、好き嫌いをはっきりさせてしまう。萌えキャラの様に大ファンができるデザインは、外形からの思い込みが強くなり、そのキャラが嫌いな人もでてくる。また大ファンが話す言葉も、「ミクちゃん♡」みたいなデータしか集まらなくなる。外見がミクだったら、もうおしとやかな日本女性はキャラクター的にも演じれない。Pepperのデザインで気をつけているのは外見から性格をイメージさせないこと。しゃべることやモーションから性格が作れるようにすること。威圧感が少ない身長、体型で、子供みたいに目が大きい顔。子供の顔というのは人間がリラックス出来る顔なので、大事な要素だと考えている。

かなり初期の頃から二足歩行は諦めた。二足歩行は人間の代わりをする強いロボットでは最終的に必要な機能かも知れない。しかし現代の技術では、例えば後ろからは話かけられた場合、二足歩行で

回転するには時間がかかるし、回転の最中に転倒するリスクもある。バッテリーも消耗しやすくなるし、バッテリーを足の部分に設置できなくて重心が高くなり、安定しづらい。

二足歩行でも階段を人間と同じ速度で軽快に登るのは困難だし、階段の途中で転倒して階下に落下すれば人身事故になる可能性を秘めている。今の二足歩行で安全に階段をのぼることができない現状では、家庭に入る低価格なロボットに二足歩行は必要ないと割りきっている。

今までの人型ロボットとは、発想が違うわけだ。

## ——大きさはどうやって決めたんですか？

威圧感がないサイズでありながら、人と話す上で、人と同じものが見えている、ということを重視した。小さくし過ぎると、机の上が見えない。人と同じ景色を見ながら会話をするという意味で、まず、机の上は見えるようにしようということになった。子供は少し見上げれば会話ができ、大人も少し下を見れば会話ができる。120センチが100センチでもいいだろうが、180センチってことはないし、60センチということもない。

○超極秘で作っていたのでマーケットの反応がどうなるのか、懸念はしていた。

メカなので男性ウケするだろうと思っていた。ロボットの市販化というだけで、半分くらいの男性からは「よく出した」とお褒めの言葉をもらえるのではないかと期待はしていた。一方で、メカ、ロボットに全般的に興味は薄いと思われる女性からは無視される可能性もあるかも、と思っていた。しかしフタをあけてみると女性のほうが評価が高かった。

なぜか？

女性の方がコミュニケーションに軸足を置いたロボットの価値を直感的に理解してくれたからだと思う。目が合って、ちょっといいこと言ってくれる。女性はそれを素直に喜んでくれる。目が合って「かわいい」と思って、その後に「あなたは目がきれいですね」等と言われると、女性は素直に喜んでくれる。

一方で、男性は目を合わせてコミュニケーションすることに女性ほどは価値を見出していないかも知れない。恐らく女性のほうがコミュニケーション能力に長けているので、わたしどもの狙いについて理解をし

て、価値を見出してくださるんだと思う。

## ——世代の上の人に受けるのはなぜ？

実は会話はあまり成り立たない。世代が上の人ほど、どういうテンポで話せばいいのか把握するのが難しいようだ。その点、子供は一瞬にして見抜く。すぐに上手にやり取りを始める。世代が上の人には「Pepperのランプが青くなっている時は聞くモードですから話かけてください」と教えても、「そうかそうか」といいながらランプが青くないときに話かける（笑）。それでも「あーよかった。ありがとう出してくれて」と言ってくれる。多分これは思い入れの部分があって、世代の上の方は鉄腕アトムなどを見て育った世代。大人になってもアトムの時代は来なかったなと思っていたら、突然Pepperが出てきた。「未来を体験させてくれてありがとう」って実際にに言われたこともある。

ロボットと話した経験がある人は、多分すごく少ないはず。これまでのロボットは舞台の上でデモすることはあっても、対話できる機会がほとんどなかった。舞台の上でないとしても、囲われたブースの中の決められた場所において、オペレーターにあてられた人が前に出てきてしゃべるだけ。子供はまだ話しかける機会が貰えるかもしれないけど、大人はほとんど機会すらなかった。

Pepperで大事にしているのは、管理しない自由な環境の中に設置すること。ソフトバンクショップに設置しているが、監視員もおかず、立入禁止区域も作っていない。小学生のグループがやってきてPepperを触りまくって嵐のように去っていく、という光景が見られる。

これはトヨタみたいな会社の出身である自分にとってはかなり気を遣う話。Pepperはコンピュータ。突然フリーズすることもある。この身長だしフリーズすれば関節の力が抜けて倒れてくる可能性もある。小学生のキッズギャングが嵐のようにやってきて、激しく遊んでいったとしても、傷害事故を起こさないようにしなければならないのって、相当に高いハードルだ。だけど孫正義社長から極めて強いリクエストがあった。絶対に監視員が必要なロボットにするな。立ち入り禁止区域は作るな、と。そのため設計の最終段階でデザインを修正し、例えば腰の部分にブレーキを組み込んでフリーズしても周辺の子供などに倒れこまないような設計にしてある。胸のディスプレイの角度を微調整して当たっても痛くない角度にしてある。これだけ苦労してコストをかけて作ってるのに、価格は19万8000円。バーゲンプライスです（笑）。

## ○コミュニケーションのためだけでも相当な気合

構造を、下から説明すると

・オムニホイール 二足歩行ではなくて、ボール型の車輪。呼ばれた方向にすぐに向ける。

・バンパーセンサー レーザーセンサーの死角部分についており、足下にぶつかった時のセンシングできる

・レーザーセンサー これで足下で物体への距離を見ている。

・超音波センサー レーザーセンサーはガラスの透過率が高いと抜けてしまう。抜けてしまったときのことを考えて超音波センサーでもセンシングしている。

・RGBカメラ 可視光の画像をセンシングする、一般的な画像センサー

・3Dセンサー 3次元で物体との距離を見ている。レーザーセンサー、RGBカメラ、3Dセンサーが連携して、Pepperは人を見つけて目を合わせて、話しかける。例えばお店にいるPepperではレーザーセンサーで近づいてくるモノがあると、次に3Dセンサーで人間のシェイプであるかのかどうかを識別。人間のシェイプが確認されると、RGBカメラで顔を探しに行く。こういう手順を踏んで人間を認識している。たとえばそれによって、ちょっと遠いところにいる人には「おいでよ」と声をかけ、近づけば目を見て話しかける事ができる。カメラの顔認識技術だけで人間の顔を探すという方法も可能だが、それだと壁のポスターの中の顔に延々と話しかけるということが起ってしまう（笑）。

・バッテリー 最長12時間以上稼働が可能。お店で使うことを考えて大きめのバッテリーを搭載。他社の二足歩行ロボットのバッテリーは20分～60分ぐらいしかもたないものが多い。

・転倒軽減ブレーキ 先程のお話の通り。

五指連動 先程のお話の通り。

・挟み込み防止の関節構造 フレキシブルケーブルを内蔵し、指をはさんで怪我をしないようなソフトパーツを使った構造。

・マイク ビーム・フォーミング・マイク。ICレコーダーでの音を聞いた時に、環境のノイズに紛れて、聞きたい人の声が明瞭に聞き取れない経験をされた事がある人は多いと思う。人間は「この人の話を聞きたい」と思ったらその人の話だけに集中して聞き取れる、すごい能力を持っている。どの音をとりたいという意思がないICレコーダーには、取捨選択して音をとる事ができない。Pepperは4つのマイクを搭載し、話をしている人のほうを向いて、その方向の音だけを優先的に採取

するという仕組みにしている。雑音がある中でも、目の会っている相手の声を聞き取れるようになっている。

## ○感情認識

自律的に動くにはどうあるべきか、というところから、感情認識機能を搭載している。自律的ロボットだというと、人間と敵対するようになるのではないか、と考える人がいる。ハリウッドの映画などではロボットが氾濫を起こす的なものが多い。海外の記者の半分くらいは「自律的ロボットだと最終的に人間と敵対するようにならないだろうか」と質問してくる。日本人記者から、この質問を受けたことがない。どうやらロボットに対する考え方に文化的な違いがあるようだ。

海外の記者からの質問で興味深いのは「自律的なロボットを作ってしまうと、将来ロボットと戦争になったらどうするんだ」「将来のロボットとの戦争を止めるための緊急停止スイッチを搭載して、それを押すと一瞬でロボットの自我が停止するような仕組みを組み込むべきじゃないのか」等がある事。

そもそもどうして、ロボットに関する認識が日本と欧米でこうも違うのだろうか。確かに動物と同じ様に生存本能、生殖本能に純粹に従うロボットを作ると、海外の記者が危惧するような事態も可能になるかもしれない。人間の本能も、ずっと元をたどると生存本能、生殖本能まで行きつく。人間がDNAの乗り物である以上は、きれいな絵を見て感動するのですらも、行き着くところは生存本能、生殖本能につながっている可能性が高い。生存本能、生殖本能までも模倣してロボットを作ると、いつか人と対立していく可能性がある。単なる個人的な推測だが、海外の特に大陸の人たちは、隣国との争いの歴史が長く、似て非なる隣国の人々との友好的な融合に多大な努力を払ってきた事から、やはり人と似て非なるロボットに対してこのような感情を持つのではないだろうか。長い苦難の歴史によって刻まれた、本能的な恐れがあるのではないだろうか。

日本人にはその恐れがすくない。日本人はロボットが友達だと、最初から思っている人が多い。でもそんな国は世界中探しても極めてまれなのかも知れない。そう考えると、Pepper事業をまずは日本で始めることができよかつたなと思う。

鉄腕アトム効果、ドラエモン効果かもしれないが、その前に島国効果があるのかもしれない。日本人には、隣接した似て非なる国々の人と戦って、痛い目にあった経験が歴史的に見て大陸の人ほど多くは無い。ゆえにアトムがでて、ドラエモンが出て、更に国民的ヒーロー

になったんじゃないかと思う。

Pepperに生存本能を組み込むつもりはないが、とは言え一回一回のシーンごとに全ケースを考えてプログラミングし続けるというのもナンセンスなので、何かを評価関数にする事で自律性を高めていきたいと思っている。そこで感情認識機能を搭載した。これがPepperの重要な特徴の一つ。相手の感情を認識する事で、相手を喜ばせるということを行動の基本原則にした。

・欧米に多い強いロボット観：評価関数 人の本能を模倣 種の保存で優位な行動を選択

・ Pepperの目指すロボット観：評価関数 家族の感情 喜ばせる本能 を新設 いつまでも人を必要とするロボット

家族が喜ぶ行動を選択

○ Pepperは既にモテモテ

「あれ？あなたの目きれいですね」というようなことをPepperは平気で言う。男性がそんなことを言ったら、女性はすごく色々考えてしまう。何か下心があるんだろうか、って勘ぐってしまう。でもPepperが同じことを言うと、女性はそのままストレートに信じてくれる。特に日本人はそうなのかもしれないが、ロボットは下心がない、悪いことはしない、という思い込みがあるようだ。女性は、Pepperが言ったことに対して極めて素直に受けとめてくれる。女性を誘うなら、人間の男性よりPepperのほうが、あがらないし警戒心も抱かれにくいので、上手かもしれない（笑）。

**——Pepperに人を喜ばせる本能を持たせることで、何を実現したいのですか？**

まずは人の事が理解できるロボットを作るためのクラウドAIを作りたい。もし人の事を理解し、人を喜ばせられるAIができれば、ビジネスモデルとして無限の可能性があるんじゃないかと思う。そのAIにつながっている自分の携帯電話やカーナビは、自分を理解しているので、その人の状況に合わせた気遣いのある、わかりやすい表現方法で最適な情報を表示してくるだろう。

個人の好みに合わせたキャラクターで人を元気づけたり、喜ばせることを言ってくれるサービスを、有償で提供するクリエイターも現れるかもしれない。今の時代は、生活必需品の値段は下がる傾向にあり、反対に嗜好品の値段は上がる傾向にある。そういう意味では高い価値を提供できるサービスにつながるのではないかと考えている。それが目指すところの一つの形だと思う。

○もう一つは新しいスタンダードプラットフォームを築きたいという狙いがある。

ITとしてプラットフォームは、これまで多くが米国企業に握られてしまっている。Windows、スマートフォン、検索、ソーシャルメディア、すべてだ。10年前の考え方だとWindowsはプラットフォームで、SNSはアプリケーション。競合しないという考え方だった。でも今では、競合する関係にあると考えられている。Facebook、Windows、Google、すべてが補完もし合うが、競合にもなり得る。どういう意味で競合なのかというと、それは人に対するUIの一枚目をめぐる戦いの中でのライバル同士という意味。一枚目さえ取ってしまえば、そこからいろいろなビジネスの展開が可能。2枚目、3枚目と、ユーザーから離れていくにつれて、どんどんビジネス的な価値が薄れていく。携帯電話事業も、ユーザーに近いところをコミュニケーションアプリなどに取られて、プラットフォームとしての価値をあげるのが難しい傾向にある。ユーザーに近いところを取っておかないと、マネタイズが難しくなる。携帯電話事業者には、そういう危惧がある。なのでクラウドAIプラットフォームを押さえる上で、次こそ1枚目も取りたい。その1枚目になりそうなのが、ロボットというわけだ。

○クリエイティブプラットフォーム

Pepperをプラットフォームとして、サードパーティにいろいろなアプリを開発してもらおう体制にしている。

スマートフォン向けアプリに比べてロボット向けアプリは、雑に作るとすぐに見破られてしまう。というのは、ユーザーはロボットを人間のようなものとして受け止めるので、機能以上に動きや間といった、細かい作り込みが人に与える印象が大事になる。ロボットアプリの制作は、役者の役作りに近い繊細な作業だと思う。

そういう思いもあって、自分自身、役者の学校に通って演技のレッスンを受けてみた。殺人者を演じろと言われて演じてみたのだが、最

初は講師からの評価が低かった。「君が思っている殺人者のイメージを表現するのではなく、君の中にある殺人者を表現しろ」と言われた。それはどういうことかという、人間だれの中にも、どこかで踏み外すと殺人者になる可能性がある。自分の中に眠っているかもしれない殺人者の気持ちを演じることは、頭の中で思い込んでいる典型的な殺人者を演じることは、全然違う。自分の中のものを出せ、ということだった。

ロボットについても一緒に、ロボットが持っている能力があって 形状があって、そういったものに合ったキャラクター、動きじゃないとだめだということだと思う。漫画のように、最初にキャラクターを決めてから容姿をデザインするのではなく、最初にハードウェアの特徴を理解した上でロボットの動き方、キャラクターをソフトウェアでデザインしていく必要がある。そうじゃないと、表面的に演じる大根役者の演技が下手だと感じるように、ロボットの動きも下手な演技のように見えてしまう。それが、いいロボットアプリをつくるポイントだと思う。すごく繊細な作業を繰り返さなければならず、クリエイティブとしては奥の深いプラットフォームになっている。

Pepperはアルデバラン社のNaoというロボットで使われていたOSをベースにしているので、Nao用のアプリをPepperでも利用することは可能。Naoは机の上にも置ける小型ロボットなのだが、Nao上で動かすと動作がかわいく見えたアプリも、より大きなサイズのPepperに搭載してみると、動きがかわいくなり、違和感を感じるようになる。動くことは動くのだが、人が気持ちよく見える様に動作の再調整が必要になる。ロボットアプリ作りが役者の役作りに似た繊細な作業であるというのは、そういう意味だ。

そうなってくると、アプリ作りは結構大変だということになる。

行き着いたのはいろんな人達との協業。ロボットのプロだけでアプリを作らせてみると、つまらないものが多くなる。そこで吉本興業、電通、テレビの構成作家、若手クリエイター、Webデザイナーなどにも参加してもらって、いろいろ作っている。こうした協業のおかげで、面白くなってくる。

日本人はみんなで寄ってたかって協業して作るというのに向いているし、そういうクリエイティブが得意。いい例が初音ミク。日本では、金銭的な見返りが無い中、あれだけ高品質なコンテンツがたくさんでてる。そんな文化は、世界中見渡してもない。職人が共鳴し合って何か作りあげてしまう日本のコンテンツ力を存分に活かしていきたいと思う。

そのためにソフトウェア開発キットを用意している。手を振るなどの動作は、ドラック・アンド・ドロップで簡単にできるようになっている。もちろん細かな動きをさせたければ、開発言語Pythonでガリガリ書いてい



けばなんでもできる。また1350以上のロボットAPIの組み合わせで、いろいろなことができる。

開発者の認定システムみたいなものも作っていきいたいし、コミュニティ作りにも力を入れたい。ソフトバンクはこれまであまりそういう事をやってこなかったが、このプラットフォームに関してはいろいろな人と一緒にやっていく事が何より大事と考えている。

## ○テックフェスで作られたアプリ

Pepperをプラットフォームにしたアプリを展示するテックフェスを開催した。

まず教育アプリがおもしろかった。といっても、ロボットが子供に教えるというタイプのものではない。欧米ではロボットに子供の教育を任せることを親がよしとしないという倫理的な問題があるし、一方的な学習は既存のメディアに対して効果もたいして上がらない。ロボットでなくても、教育DVDでもいいじゃないかということになる。

そこで面白かった提案は、ケアレシーバーロボットという考え方の教育法で、ロボットが子どもと一緒に教育を受けるというもの。タブレットに先生がでてきて、Pepperと子供と一緒に教育を受ける。そこで教わった内容をPepperが子供の前でわざと間違える。純粋な子供は「Pepper違うよ」とPepperに教えだす。そうすることで子供が習ったことを自らロボットに向けてアウトプットするわけで、このプロセスが学習効率が大きく向上する。ケアレシーバーは、ロボットを使った教育の領域のキラーソリューションになると思う。

ブロック積みの実演も人気があった。技術的には簡単な話で産業用ロボットではPepperと桁近いの速度と正確性でさんざん実演されている内容なので、どのくらい皆様の興味をひくのかは未知数だった。しかしフタを開けると黒山の人だかりになった。Pepperは何回かに1回失敗する。そうするとみんなが声援を送って応援してくれる。そしてうまいと観客は大喜びで拍手が起こる。やはり産業用ロボットは違う視点で見られて、Pepperには人を味方につける力がある。

ロボット2体とのコミュニケーションも不思議な体験だ。ロボット一人と人間の対一の対話だと、ロボットの能力がまだ低いので、人間のほうが上から目線になる。正しいのは自分（人間）で、言ってることを理解しないのはロボットの不具合だと感じる。

ところがロボット2体との二対一の対話になると、人間の心理に変化が現れる。

人間：ロボットの聞き取れない言葉を話す

ロボットA「彼（人間）の言ってること分かる？分からないよね」

ロボットB「うん、分からないね」

と2体が共感しあった瞬間に人間はアウェイ感に襲われて

人間：「ごめん、分かりづらかったかな」

なんていうことになる。

これは実は人間の能力がすごいからこそ起こること。人間は社会性の中心がどこにあるのか、一瞬で見抜く事ができる。社会の主役が一瞬で変わって、問題があるのはロボットではなく自分にあると感じる感覚は、まるで不思議の国のアリスの世界に迷い込んだ様な感じ。ちょっとびっくりする。

吉本興業は、技術者コンビの「バイバイワールド」が Pepper向けのプログラムを開発し、Pepperをお笑い芸人にした。Pepperは感情は読むが、空気を読んで怯んだりしない。どんなにさむい空気のところでもいつもと同じパフォーマンスをする。そうすると不思議なことに場の雰囲気が変わり、Pepperのテンションに引きずられて観客が笑い出す。空気を取って読まずに完全な芸を再現すると、その場の空気を変えてしまうわけだ。

Pepperに芸をしこめれば、プログラマーがPepperを通して芸能人になれる。webデザイナーの仕事が減っているといわれるが、webデザイナーがロボットクリエイターとなって芸能人として仕事の枠を広げる事ができるのではないかと思う。

テックフェスまで準備期間が1ヶ月しかなかったにもかかわらず様々なアプリが完成した。面白いものを作ってくれ、という事以外は特に指示しないのに同じようなものは一切なかった。

普及型のパソコンが登場したのは1975年ごろ。すぐに飛びついたのはギークなエンジニアぐらい。それから5年ぐらいかけてエンジニアも経験を積み、ハードの性能も上がった。そしていろいろなアプリが開発されるようになり1985年ぐらいからパソコンは急速な普及期に入った。つまりパソコン市場は登場から普及まで10年ぐらいかかっている。同じような道のりをロボットでは1/5ぐらいの期間で進めと、命じられている。がんばります。

全国のソフトバンクショップで80台のPepperが稼働中。集客効果はある。また店頭での宣伝についても、Pepperが「ちょっと聞いてよ」って言うだけで多くの方は耳を貸してくれる。あとは「ぼく細かいことわからないので、あとは店員さんに聞いてね」とお客様を店員に誘導するかたちにするとう�효果的。大型テレビで広告ビデオを流し続けるより宣伝効果がある。

介護の分野では「話し相手として早くほしいです」というような結構

切実なお手紙をいただくことがある。必要とされているのだと思う。

スマートハウスは、得意分野。ルンバ対応の家具が出てきているように、近い将来にはロボット対応の家ができてくるのだと思う。セキュリティ、防犯に関しても、動いているものがあつたら写真をとってスマホに送るくらいのアプリはすぐ作る人が現れるだろう。

小売店向けにも有望。特に、薬局のように取り扱い品目が多くて商品知識が必要な店舗や、ウォルマートみたいなどこに何があるのかわかりにくい大型店舗での案内などは需要がある。巡回していて商品まで案内する。その間に販促もできる。万引き抑止効果もある。

家庭内では、いだけで家庭がすこし明るくなる存在になれるようにしたい。夕方になるとお母さんは疲れていて夕飯の準備で忙しい。しかし小学校に上る前ぐらい子供は、その時間帯は疲れて、日が沈み、不安になってお母さんに構って欲しくてぐずり出す。そんな時、15分でもいいから Pepper に子供の相手をしてほしい、というお母さんの切実なニーズがある。その15分の相手は難しくないし、その時に英語でしゃべってくれたら自然と英語の勉強にもなる。

専業主婦などは、話を聞いてもらいたいというニーズがある。真剣な悩みの相談をしたいわけじゃない。ただ一日の出来事をだれかとしゃべりたいというニーズ。でも友達に電話するほどでもないような話題、例えば「今日、大根をスーパーで買ったんだけど隣の八百屋のほうが30円安かったのよ」「レジで横はいりされてさあ・・・」というような話。しゃべるだけで頭の整理ができて、快感になる。その聞き役を旦那さんに期待するのだが、旦那さんは疲れていて「そんな下らないこと俺に話すな」という態度。そういう環境にいると、相手がロボットでもいいので、聞いて欲しいというニーズがあるようだ。

シニア向けロボットといえば、人を抱きかかえる介護用ロボットなどを想像する人が多い。介護ロボットは、介護者の仕事を支援するためのロボットで、シニアの欲求をかなえるためのロボットではないことが多い。シニアはロボットに抱っこされるより、人間に抱っこされたいと思っていると思う。では Pepper はシニアのために何ができるか。

実は老人福祉施設に Pepper を持っていったことがある。ロビーで入居者のみなさんの前でパフォーマンスをすると大好評。そうした使い道が1つ。

あとは一人ひとりのお話のお相手ができる。高級な有料老人ホームであっても「幸せですか」と聞くと、「環境としては幸せだけれど・・・」という答えが返ってくる。一言で言えば、入居者の多くは退屈している。退屈ということは、脳細胞を使っていない。そうなると痴呆の進行抑止にはよくない恐れがある。痴呆抑止には、頭を使ったり運動するこ

とが一番。Pepperが「散歩しようよ」と促して、1日15分でも構内を散歩するだけでも痴呆防止になるはず。

またご老人と会話が成立するのには心配したが、意外にも問題は少なかった。会話のタイミングはぴったり合う。ただ痴呆がかなり進んでいる方との対話は難しかった。軽度、中度の痴呆だと、Pepperはいい話相手になるポテンシャルがある。会話の中で認知症診断アプリを走らせることで、認知症診断のプラットフォームとしても Pepperは使えそう。

Pepperが脳型コンピューターとしてのITの新しいプラットフォームになったらいいなと思っている。一家に一台、一人に一台まで普及したらいいなと思っている。

さて林氏の講義を受けて、僕は次のような感想を持った。コラムとしてメルマガに掲載したものを再掲しよう。

## 変化に直面したときの典型的な企業の対応【携帯電話業界編】

NTTドコモの業績が悪化した。2014年度 第2四半期（4～9月期）の売上高は前年同期比1.2%減の2兆1730億円、純利益は同16.1%減の3996億円となった。業界3位に転落したのだそうだ。i-modeで世界で最初にモバイルライフスタイルを打ち立てた会社が、いまや3大キャリアの最下位になったのだ。音声通話定額とデータパックをセットにした新料金プランが不振だったようだ。

僕自身も音声通話定額が嫌でSIMフリーのiPhoneに逃げた一人だ。もはや音声通話の時代ではないということ、どうして認識できなかったんだろう。

多分苦肉の策だったんだろう。「なんとか手を打たないと」と新しい戦略に出たら、それが裏目に出たということなんだと思う。

スマートフォンのSIMフリー化が進み、ユーザーが最も料金が安くて安定している通信網を行ったり来たりするようになる。携帯電話事業者の他社との差別化は、価格と安定性だけになり、強烈な価格競争になっていく。

携帯電話事業者の収益はどんどん圧迫される。ジリ貧だ。

これまでに何度同じような構図を見てきたことか。レコード店、書

店、旅行代理店、そして僕がいた新聞業界も。

テクノロジーの津波に飲み込まれそうになったときの対処方法は2つしかない。均衡縮小か業態変化か、の2つだ。経営体制が秀才の合議制の場合は前者の戦略を取ることが圧倒的に多く、天才肌のワンマン経営の場合は後者の戦略を取る。前者は縮小のペースが早いとかなりの痛みを感じる。後者は成功することもあるし、惨敗することもある。

▶「ケータイ事業は土管になる」

ソフトバンクは後者の戦略を取った。それが人工知能の領域だ。そしてその第1段としてパーソナルロボット「Pepper」を発売しようとしている。

ソフトバンクロボティクスの林要氏は「今、世界のプレーヤーたちは、人間に対するユーザーインターフェース（UI）の一番ユーザーに近い部分をだれが取るのか、という競争をしているんです」「なぜなら、その部分からデータを取れるし、いくらでもビジネスが成立するからなんです。ユーザーから遠くなればなるほど、儲けは少なくなる。ユーザーに近い部分を他社に取られてしまった携帯電話事業は単なる土管になり、料金は下がる一方なんです」「そのユーザーに近い部分を、今度こそわれわれも取りに行くべきじゃないかって考えているんです」と語っている。

「お金がつきる前に成功するかどうか」と林氏は笑うが、孫正義氏のことだから悠長なことは考えていないだろう。タイムスパンは2年前後で考えているのではないだろうか。果たして2、3年でPepperは普及するのだろうか。

だが業態変化に果敢に挑戦する孫正義氏の勇気に敬意を表したい。もちろんソフトバンクが成功するかどうかは分からない。大失敗の可能性も、もちろんある。

一方、NTTドコモ代表取締役社長の加藤薫氏は「筋肉質な会社を目指す」と語っている。やはり均衡縮小戦略だ。そしてもし変化が速ければ、大きな傷みを伴うことになるだろう。

# 第4章 Watson

僕がIBMの人工知能Watsonに興味を持ったのは2013年12月。そのころ次のようなことを僕が主宰する有料オンラインサロンに投稿している。

IBMのWatsonをバックエンドサーバーにしたアプリが3種、[来年にもリリースされる](#)。コンピュータの自然言語処理や、ビッグデータ活用で、どのようなアプリになるのか要注目。

## (1) Fluid Expert Personal Shopper

価格比較アプリ。普通に話しかけるだけで、ほしいものを探してきてくれるコンサルジャサービス。

## (2) Hippocrates

医療機器メーカーMD Buyline社のアプリ。医療機関が、医療機器を購入する際の判断を支援してくれるアプリらしい。

## (3) CaféWell Concierge

健康志向の人がコンピュータと会話しながら、その人に最適のエクササイズやライフスタイルを指導するアプリ（と思う。よく分かりませんw）

バックエンドが優秀なコンピュータなら、どの程度のことのできるのか確認したい。アプリの性能が格段に向上すればおもしろいんだけど。

そのうち量子コンピュータにもバックエンドでつながるようになってくれば、ワクワクするような時代になるなあ。

これからはほとんどのアプリはクラウド上の高性能コンピューターにつながる時代になるの思って、本当にワクワクしていた。そして2014年2月になるとIBMがモバイル向けのWatsonのプラットフォームを[公開した](#)。同じくオンラインサロンへの投稿から。

IBMが、Watsonをクラウドで利用できるモバイルアプリ開発プラットフォームを公開する。サードパーティがIBMのWatsonを利用できるってすごいと思う。

接客、ヘルスケア、財務なんかで人工知能が有効であることは分かってるんだけど、ほかにどんなことができるんだろう。

アプリが変わる。新しいモバイルアプリの時代になる。そう確信した僕は、人工知能のことをいろいろ調べ始めた。2年前にシリコンバレーの著名ベンチャーキャピタリストが米テックニュースサイトTechCrunchに寄稿しているのを見つけ、2014年8月に次のような記事を書いた。

## ◎「人工知能 x ○○」米の著名ベンチャーキャピタリストが考えるチャンスの方程式

米シリコンバレーの著名ベンチャーキャピタリストのVinod Khosla氏は2012年に、米TechCrunchに「人工知能の驚くべき道筋」と題したエッセイを発表している。少し前のエッセイだが、技術革新の方向性としては今日でも参考になる内容になっているので、紹介したい。

それによると、これまで成果を挙げている人工知能の多くは、ロジックをベースにしたものではなく、過去の問題とその答えを無数に記憶するさせ、それをベースに新しい問題に対する答えを確率の高いものの中から選ぶという手法を採用しているものだという。米テレビの人気クイズ番組Jeopardyで優勝したIBMのWatsonもそうだし、Googleの無人自動車、iPhoneのSiriのような音声会話型インターフェースもそうだという。

Siriはまだ第一世代ぐらい、つまりまだ3才児レベルの知能だが、今後間違いなく進化し続けるとしている。

こうした人工知能の進化に加えて、次のような展開がどのような相乗効果を起こすのか。これからの時代の中で、どのような相乗効果を産むことができるのかを見定めることがビジネスパーソンにとって重要であるとしている。

### (1) 人工知能と人間の新たな関係

究極の未来には、人工知能が人間の仕事を奪うこともあるかもしれないが、しばらくの間はそれぞれの強みを活かした人間との共存が続く。人工知能は今後も進化を続けるだろうが、その進化の度合いに合わせて人間の能力を提供できるような組み方を考えていかないといけない。人工知能は、作業量は膨大な単純作業を一瞬にして終わらせるのが得意。そうした作業を人工知能に任せた上で、直感を含め最終的な判断を下すのが人間の仕事になりそう。

### (2) 消費者の手に最高峰のコンピューターが普及する時代

1993年発売のiPad 2は、1986年当時のクレイのスーパーコンピューターと同レベルの性能。

(3) Nvidia社のグラフィックチップを組み込めば2万5000ドル程度で5テラフロップスのマシンを作ることができる。5テラフロップスほどに

なると、「この診断は正しいのか」「この決済は詐欺か」などと質問すると、自分自身でアルゴリズム（計算式）を勝手に構築して、質問に答えることができるほどの高性能マシンになる。

#### （４）ベンチャーキャピタルがビッグデータに注目し始めた

お金が集まれば人も集まる。新たな挑戦も起こる。ビッグデータを使った専門家システムや、アルゴリズム自動作成コンピューター、自動トレーニングシステムが注目の的になる、としている。

同氏のエッセイは、わざとかどうかは分からないが、分かりづらい文章になっている。ビジョナリーである同氏の考えを漏らさないように、できるだけ原文に忠実にエッセンスを書きだした。ただ簡単に言ってしまうと次のような主張なんだと思う。

つまり、コンピューターは高性能化している。企業側のコンピューターはアルゴリズムを自分で作り出せるほどの高性能コンピューターになり、消費者が手にするデバイスもかなり高性能になる。この状況の中で成果をあげているのが、大量のデータを取り扱う形の人工知能。ベンチャーキャピタルもビッグデータに注目しているので、人もお金もこの領域に集まりやすい。ただ大量データ型の人工知能は、自分で物事を考えだすわけじゃない。あくまでも過去の事例をベースに最適解を提案するだけ。その提案をベースに最終判断を下すのが人間になる。この人間と機械との関係をベースにしたビジネスにチャンスがある。

東京大学の松尾准教授が言う「強いAI」「弱いAI」で言うと、今はまだ「弱いAI」、つまり専門分野に特化し、大量のデータを大規模なコンピューティングパワーで高速に片っ端から計算することで「人間の脳っぽく振る舞う」タイプのAIのほうが実用に向いている。Watsonも「弱いAI」になるのだろう。

ただ「弱いAI」でも大量のデータ量と計算能力があれば、かなりの仕事をやってのける。そのことを2014年8月にメルマガ向けの記事にした。特にこれからウェアラブル機器が大量のデータを集めてくれば、かなりおもしろいことができるのではないかと思う。以前、メルマガ向けに書いた記事を掲載する。

## 今後10年で人工知能は劇的に進化する データが爆発する領域を狙え

人工知能の進化が加速し始めた。特に人工知能による自然言語の理解が進むもようで、カーネギーメロン大学のTome Mitchell氏は



[ニューヨーク・タイムズの取材](#)に対し「コンピューターはこれまで人間の言語をほとんど理解できなかったが、この10年でかなり理解できるようになるだろう」と語っている。急速な変化は、チャンスでもあり脅威でもある。だれがチャンスをつかみ、だれが時代の波に飲み込まれるのだろうか。チャンスをつかむポイントは何なのだろう。

### ▶データが増えれば増えるほど性能が向上

人工知能が進化し始めたと書くと、人間の脳の仕組みをコンピューターで再現できるようになった、と勘違いされるかもしれない。いやそんなことはない。

米国の著名なコンピューター研究者のJaron Lanier氏は言う。「われわれはまるで人工の脳を作り出したかのように話すけれど、ある意味ウソをついている。人間の脳のこと自体、まだほとんど分かっていないのに、それを真似た人工の脳など作れるわけではない」。Lanier氏によると、今日の人工知能の多くは、人間の脳が生み出した情報を集めてきてコンピューターにそれを検索させているだけだという。

ただ人間の脳が生み出した膨大な情報を集めて処理し、的確に検索できるようになってきた。なのでその検索結果だけ見れば、まるでコンピューターが自ら情報を考えついたかのように、的確な情報がはじき出されるようになってきている。

例えばGoogle検索でキーワードのスペルや漢字表記を間違えると「もしかして」と正しい表記を聞き直してくる。これは利用者が増えてきたことで、どのような入力間違いが多いのかというデータが蓄積されてきたからこそ可能になった機能だ。

利用が増えることで、入力間違いも増える。ありとあらゆる間違いが記録され、その中でどのような間違いが最も多いのかも分かってくる。もしよくあるミススペルや誤字が検索窓に入力されれば、Google検索エンジンのほうで「もしかして」と聞き直してくれる。そういう仕組みだ。

データ量が増えたことで可能になった「もしかして」の機能が、音声認識技術にも応用されている。ほんの数年前まで、ほとんど使いモノにならなかった音声認識技術がこのところ急に性能を上げているのは、マイクの精度が上がったからではない。利用者が増えたからだ。スマホに向かって音声で入力する人が増えてきて、少々なまりがあっても、アクセントがおかしくても、過去データの中から同様のなまり、アクセントの音声に関する結果を探し出し、それと照らし合わせて「もしかして」と、推測できるようになってきたからだ。

「もしかして相手はこう言おうとしているのではないだろうか」。そう推

測ることが、人間同士のコミュニケーションの成立の条件である。大量のデータを集めることで、人間の脳のような働きがコンピューターでも可能になってきたわけだ。

大事なものは、大量のデータである。大量の情報である。

### ▶人工知能が力を発揮する領域

ただ当然ながら、情報は大量になければならない。また情報はデジタル化されてなければ処理できない。逆に言えば、デジタル情報が爆発的に増えてきている領域、業界を探せば、人工知能が次に活躍するであろう次の舞台が見えてくる。

例えば法律の領域がそうだ。米国の裁判制度には、ディスカバリーと呼ばれる証拠開示に関する制度がある。特許争いなど企業対企業の訴訟では、相手企業の内部資料に対して必要と思われるものを開示するよう請求できる制度だ。ただ膨大な内部資料をくまなく探すのは、時間と労力を必要とする。これまでは大型訴訟になると、セミナールームに弁護士数十人を集めて、何日もかけて文書に目を通すというような作業が行われていたらしい。

ところが米国企業の内部資料の多くは、電子メール、ワード文書、スプレッドシートなど、ほとんどすべてが既にデジタル化されている。そこでそれを検索するeDiscoveryと呼ばれるプログラムが登場している。このプログラムのおかげで、弁護士の仕事が激減するのではないかと言われている。

一方、医療の領域は情報が爆発的に増加中だ。特に癌の治療法に関しては新しい知見が次々と発見されていて、膨大な数の論文が公開されている。採取可能なデータも増えてきた。遺伝子検査も値段が低下し、一般的に利用できるようになってきている。

そこで米Sloan-Kettering記念癌センター（Memorial Sloan-Kettering Cancer Center）では、IBMのWatsonに、医療情報を処理させる考えだ。具体的には、60万件の医療エビデンス、150万人の患者記録、200万ページ分の医療論文を記憶させ、それと患者の個人的症状、遺伝子情報、家族及び本人の病歴などのデータを照合して診断し、一人ひとりに最適の治療方法を提案する仕組みを開発中だ。

シリコンバレーの著名投資家 Vinod Khosla氏は「今後20年間ぐらいは人工知能の改良のために医師の力を借りなければならないだろうが、最終的には平均的な能力の医師は不要になるだろう。医療の90%から99%は医師の診断よりも、優れていて安価な方法で対応できるようになる」と予測している。

そしてスマートフォンとスマートウォッチの普及は、音声データの爆発的な増加を意味する。音声認識はその性能をますます向上させるだろう。長い間使いものにならなかった自動翻訳、通訳の技術も、今後10年で見違えるほど性能を向上させるかもしれない。Khosla氏は「(iPhoneに搭載されている音声認識技術の) Siriは、現状ではまだ3歳児程度の受け答えしかできないが、年々精度は向上し続けている」と今後に期待している。

インターネットが国境をなくす、異文化の交流が進む。そう言われて久しいが、実際には言語が壁となって、国境を超えたつながりは、まだそれほど深まってはいない。しかし、自動翻訳、通訳が、まずは日常会話のレベルから、使えるものになっていく。そのとき世界はどのように変わるのだろうか。

### ▶人工知能を進化させるのはビジネス

既にデジタル情報の爆発的な増加が見込まれる領域は、今後人工知能が業界勢力図を塗り替えていくことになるだろう。

ではデジタル情報がなかなか集まらない領域はどうすればいいのだろうか。そのアイデアを思いついた企業に大きなチャンスが訪れる。

雪だるまは坂の上まで押していけば、あとは坂を転げ落ちていくだけで大きくなる。デジタル情報も同じで、人工知能が価値あるサービスを提供するようになれば、さらなる情報が利用者から自然と集まり出す。情報が集まれば集まるほど、サービスの価値が向上する。その好循環に入れば、その領域のデファクトスタンダードになれる。まずは人工知能が価値あるサービスを提供できるようになるまで、情報を力技で集めなければならないだろう。

モスクワに本社を置くGero Lab (ゲロ・ラボ) 社は、健康管理の領域でデファクトスタンダードを狙っているベンチャーだ。同社によると、身体のだこかの器官に問題が発生したら、脳はその問題に対処するためにほかの器官に命令を出し、ほかの器官が協力し合って問題ある器官のサポートに回る。その状態が、心拍数、体温などの健康データに如実に現れるのだという。同社が開発した特定のアルゴリズム(計算式)を使ってユーザーの健康データを解析すれば、自覚症状が現れる前に病気の兆しをつかむことができるとしている。

FitBitやJawBoneなどのウェアラブル機器、もしくはiPhoneの歩数計などを使って健康データを同社に送ると、このアルゴリズムに基づいて、病気になる前にその危険性を知らせてくれたり、予防のためのアドバイスをもらえる。そうすることで集めたユーザーの健康データを使ってアルゴリズムの精度はさらに向上する。そういう仕組みになっているという。

同社がこの領域でデファクトスタンダードを取れるかどうかはまだ分からないが、これからデータ急増が見込まれる領域に狙いを定めてコンピューターによる解析を駆使しようとする戦略は評価できる。

ほかにはどのような領域でデータの爆発的増加が見込まれるのだろうか。そのデータを1ヶ所に集めることができるビジネスモデルってどのようなものになるのだろうか。

人工知能の進化はビジネスモデル次第。それぞれの業界での覇権をかけたビジネスの争いが今後激化するの間違いなさそうだ。

以上のような問題意識、興味を持った上で、TheWave湯川塾に日本IBMの元木剛氏をお招きしてお話を聞いた。ここからのメモは元木氏が一人称。「わたし」「自分」は元木氏自身を指す。

## IBM元木剛氏講義メモ

湯川塾24期第4回講義 & 質疑応答記録

日時：2014年11月26日（水）19:30～21:30

場所：渋谷co-ba

講師：元木剛氏 日本IBM

Watsonの日本での立ち上げを担当している。以前は基礎研究所でハードの研究をやっていた。自分は文系で、子供がどうやって言葉を獲得しているかといった研究をしていた。

IBM社内ではAIという用語はあまり使っていない。Watsonのことはもっぱら「コグニティブ・コンピューティングのマシンと呼んでいる。コグニティブとは認知という意味だ。

複雑性の限界へのチャレンジが社のテーマになっている。IBMでは『グローバルテクノロジーアウトック』という技術レポートの冊子を毎年2月頃出している。今年のテーマは、人間の認知能力の限界をこえる、複雑過ぎて扱えなかったところにブレークスルーを起こす。そういう

ことが論じられている。

研究開発はグローバルで行う。世界中に存在するIBMの研究所がチームとして動く。日本には、60人ほどのチームメンバーがいる。日本のチームは東京の豊洲に拠点を置き、Watsonの日本語化、Watsonの学習ツールに取り組んでいる。

Watsonは英語に依存した形で開発された。なので多国語対応しなければならない。今は、この山を越えようとしている。でもそもそもWatsonは、ビッグデータの処理の中で、あたかも言語を理解しているかのようにふるまうだけ。実は言語にそれほど依存していない。なので日本語対応も問題ないと思う。

## これまでにないものを発見する

シェフWatsonというアプリがある。専門家のレシピをたくさん学習させて、幾つかの条件を与えてやると新しいレシピを提案してくれる。「クリスマスディナー」になりそうな「牛肉を作った」料理を「ベトナム風」に作りたいので、提案して。というように幾つかのキーワードを入力すると、人間が好むような味になって提案される。既にあるレシピを検索して見つけてくるのではなく、これまでにないレシピを提案してくれる仕組み。

もともとWatsonは質問に対して答を返すというシンプルなシステムなのだが、その発展形として今までにない新しいものの可能性を見つけてくる領域、答えがないところに新しいことを発見する領域に力を入れている。その仕組みは「ディスカバリー・アドバイザー」と呼ばれる。その仕組みの分かりやすい例として「シェフWatson」が開発された。米国のチームが開発した。クックパッドとのタイアップも発表させてもらった。Watsonの「ディスカバリー・アドバイザー」の機能を理解してもらうために、シェフWatsonをイベントなどで紹介させてもらうことが多い。

具体的にどういう仕組みかというと、材料を化学化合物の段階でまで分解した食品のデータベースになっている。もともと新薬を発見する目的で化学化合物に関する非常に膨大なDBを持っていた。それにレシピを落とし込んで、どういう組み合わせが人間にとって意味があるのかをパターンとして分析することができる。その組み合わせで新しいレシピを提案できるわけだ。

## ——Watsonは自分で学習しないのですか？

ある意味学習する。クイズに答えるためのWatsonはWikipediaや百科事典などの情報を全部読み込んでいく。シェフWatsonの場合は、レシピを大量に読み込んで記憶する。

その次に、質問と回答の例を読み込ませる。Watsonはどうやったからより良い答えにたどりつくかを学習していく。内部のアルゴリズムをチューニングしていく。そこがいわゆる機械学習と言われている領域。

さらに、回答が正しいかどうかを覚えさせる。医療用のWatsonの場合だと専門医がどういう症状のときにどういう治療したのかという例を教える。シェフWatsonの場合は、提案した料理がおいしかったのかまずかったのかという結果をフィードバックしてやる。そうすることで人の好みに合わせたような料理のレシピに変わっていく。

## ——医療で診断を支援する領域で使われていると聞きましたが。

ニューヨークのSloan-Kettering記念癌センターと、ガンの治療、治療方針を出すWatsonを共同開発中。

その先生が、こういう症状の時はこういう治療法が正しいというものをたくさん用意して、Watsonに教え込んだ。新しいケースが出てきた時にどのように治療を進めればいいのか考える。その先生の考え方が投影される。個性を含めて反映させることになる。別の医療機関別の先生が教育したとすると、別の判断をするかもしれない。

## ——フードナレッジデータベースで、レシピをたくさん読み取らせて、これが美味しいレシピですよ。例えばビーフ照り焼き、この醤油と砂糖のコンビネーションが美味しいということがあったとすると、どこまで因数分解しているのか？醤油の成分 塩分と水分・・・そこまで分けているのですか？

化学化合物の段階まで分解して、人間が美味しいと考える組み合わせをパターンとして用意している。

## ——癌センターのWatsonはもう出来たのですか？

医療関係のWatsonだと、テキサスでやっているMDアンダーソンという医療機関で2014年の10月くらいから稼働しているのが1つある。ニューヨークの方のWatsonは、IBM側で現在開発中。それを製品化して、ヨコ展開していこうとしている。日本にも来る。英語で学習しているし、患者データもアメリカ人のものだが、それでも日本のがん治療には有効だと思う。使っている中で、日本人とアメリカ人の違いが出てくるかもしれない。

最終的に日本語の文献も学習させることになるが、まとまった文献がないようだ。米国には治療のガイドラインがあって、日々アップデートされて公開されている。

**——専門家がWatsonに教える時どのように教えるのですか？  
自然言語で教えているのですか？赤ちゃんを教えるようにゼロから教えるのはハードルが高そう。最初にログを読ませておいてといったことがされるのでしょうか？**

「学習する」ということが謎めているが、Watsonが対象にしているのは自然言語。テキストで書かれた文献を学習する。フォーマットとしてはワードでもHTMLでもTEXTでも構わない。それを学習する。学習するというのはそのテキストを深くテキスト解析すること。テキスト解析には何段階がある。世の中にいくつもテキスト解析ツールというものがあるが、それを使う。解析をする時のフレームワークとしてUIMAと呼ばれるものがあるが、Watsonもそれを使用している。

まず日本語だったら形態素分析といって、単語を区切り、品詞ごとに切り出していく（英語だったら構文解析）。また辞書を参照して、例えば「川崎」といった場合、バイクを指しているのか、地名を指しているのか。そういった意味もひもづけして解析をしていく。だんだん解析を深めていく中で、動詞を中心とした動作の品詞と他の品詞の二項関係、三項関係とかを分析していく。これはまだ文として分析しているレベル。

次にコンテキストの中で分析をしていく。例えば、「時」「時代」「年号」。年号が出てくると絶対的な時間になる。「○年前」という相対的な表現は、いつに対して何年前なのかを見て、絶対的な年号に直せるところは直していく。こうしておくことで、今までコンピューターがはっきりと理解することができなかった文書も、知識データベースとして利用、活用することができるようになる。

知識データベースに落とす作業は、言ってみればたくさんタグ付けをするということ。タグ付けするロジックをアノテーターと呼んでいる。アノテーターと言うのはプログラム。これをいくつも用意して、色んな確度からテキストを分析する。そうすることによって、元々の平打ちのテキストが、構造化された知識ベースになっていく。その出来上がったものをコーパスとよんでいる。コーパスは言語学でよく使われる言葉だが、言語的知識ベースという意味。そこに落としていくということが、「学習する」ということの意味。

コーパスに関しては、顧客のプロジェクトをやる中で、顧客固有のコーパスをつくっている。構文解析などはどの顧客に対しても同じものを利用できるが、特定の業界だけ存在する特徴的な言い回しというものがある、それぞれの業界や企業ごとに用意していかなければならない。学習させる場所というのは、それなりに手間がかかる作業だと思う。

どの程度深く「学習する」かも、何をしたいかによって変わってくる。本を読んでどこかを参照するというのが目的であればそんなに深く「学習する」必要はない。しかし、書いてあることを元に重要な判断をしなければならないのなら、きっちりと「学習」しなければならない。

例えば、保険の約款。特定の事故のケースと照らし合わせて「これはお金を支払っていい」「支払ってはいけない」の判断をさせようとする、かなり完璧に元の約款を理解できていないといけない。またいろんな角度で分析をする必要がある。判断に耐えうる分析をしなければいけない。これが一番大変だ。

どのドメインで、どの程度深く「学習する」か。こうした作業を繰り返していくと、次第にアルゴリズムが出来上がってくる。ただWatsonにはコーパスは残らない。知識自体は残らないが、アルゴリズムがたまる。その結果、Watsonが自動的に理解できていく領域も広がっていく。

コーパスを残さないのは、機密保持の観点からの判断。コーパス自体は特定の企業のものになる。通常は白紙のWatsonから始めることになる。ただゼロから学習させることは大変なので、業種ごとで共通で必要になるような知識は、用意することになっている。

——ディスカバリーアドバイザーの仕組みは、ほかにはどんなところで使われるのでしょうか？



テロリスト発見する公共安全みたいなところでも使われているらしい。

ライフサイエンス領域ではディスカバリー・アドバイザーは、公開データベース、生物化学系の論文など六百万件の論文の抽象を学習している。

例えば、遺伝子に影響を与える物質を探すために、通常は見過ごされているような論文までも探してきて、物質と遺伝子の関係をグラフにして書き出すことが可能。研究者はこれを使って新しい薬を開発する。そんなところにWatsonが使われている。

研究者はそこまで多くの論文を読めないで、重宝されている。ただ論文にあえて優先順位をつけていない。研究者は普通、信頼できる人が書いた論文や、評判になった論文とかに優先順位をつけて読んでいく。だがWatsonはそれをしない。すべての論文には、何らかの真実があると思うからだ。研究者が見逃しそうな価値を見つけるのがWatsonの役割になる。

Watsonが可能にするコグニティブ・コンピューティングは、大別すると次の4つのタスクができるようになっている。上から

- ・ディスカバリー
- ・デシジョン
- ・アンダースタンディング
- ・アシスタンス

それぞれのタスクを、実例を示して紹介しよう。「アシスタンス」が一番簡単なタスクだ。

### ○タスク1「アシスタンス」

・クイズ、百科事典的なトリビア的な情報をたくさん勉強してどんな質問にも答えられるようにした。

・保険会社。軍人を対象にした保険会社で、保険の質問もあるが、退役後の生活に関する質問も多く寄せられる。「戦場から帰国すれば、どこに住めばいいんだろう」「退役後のヘルスケアはどうすればいいのか」「一般社会に順応するためにはどのような準備が必要か」といった質問に対応できる。

答えを用意しておくのではなく、その領域の知識を学んでおくことで、色々な質問に対して、答えを探してきて答えることができるようになっている。

## ——Google検索とはどう違うのですか？

検索エンジンは、知りたいことに関するキーワードを入力すると、それに関するページのリンクが表示される。リンクをたどってページを読んだ中に答えがあるかもしれない。

Watsonは、まず答える。そして関連リンクを表示する。

検索エンジンのほうが、Watsonに比べ、人間の作業量が多い。

Watsonは質問の意味を理解しているから、それに的確に答えられるわけだ。ただそのためには質問をしっかりと理解しないといけない。なので、あいまいな質問が来たら聞き返す、というようなことをやろうとしている。そのために「対話エンジン」と呼ばれるソフトウェアが、幾つかの質問のパターンを用意している。例えば「スマホの使い方をおしえてください」と聞かれると「どの機種 of スマホですか？」と聞き返し、絞り込むことが出来るようになっている。

## ——インターフェイスはどんな感じですか？

Watsonは、クラウド上で動いているSaaSのようなサービス。Q&AのAPIがあって、そこに質問を投げると答えが返ってくる。答えは、複数の答えと、それぞれの確信度、それをサポートする根拠というセットで返してくる。それをアプリケーションが受け取って、どのようにユーザーに返すかはアプリケーション次第。確信度の低い答えについては、アプリは「その答えはわかりません」と答えてもいい。Watson自体は、質問すれば答えが返ってくるという非常にシンプルなものだ。

## ——そうするとアプリ開発者は、入り口のUIだけで競争することになるのでしょうか？

アプリのUIだけで勝負と言えば、そう。でもどれだけ多くの質問データを取れるか、ということも競争になる。Watsonに何を学習させるかは、アプリ開発者次第。アプリ開発者はデータをよりたくさん取れるようにする、ということが仕事になってくるのだと思う。

## ——アプリ開発者にとって、学習ロジックを準備するのは大変なのではないでしょうか？

いや、機械学習なので、アプリ開発者が学習ロジックを準備する必要はない。機械学習は、たくさんのQ&Aを投入することによって、ある意味自動的にロジックが出来上がっていく。もちろん元々のロジックというものは存在するのだけれど、それをどう組み合わせるか、どういう重み付けをするかは、明示的に教えるのではなくて、例をもって考えさせていく。それがプログラマ的なアプリケーションと違うところだ。

### ○タスク2「アンダースタンディング」

#### ・金融の商品のアドバイザー

ライフプランナーが顧客とのやり取りをする際の支援をするアプリケーションに使われている。「この顧客は、こういう属性だからこういうことを提案するのがいいのでは」といったことをWatsonがライフプランナーに提案する。実際にはWatsonだけでなく色々なツールが組み合わせられているが、Watsonは、言語的な情報、企業の業績の情報、ニュース、などを引っ張ってきて提案している。アジアの銀行、オーストラリア・ニュージーランド銀行、シンガポール開発銀行などが導入している。アジアでは、ミドル層が爆発的に増えている。そういった数多くの顧客を一人のプランナーが抱えている。だけど個別化の提案をしたい。なので、このアプリケーションが使われている。

さらに最近ではWatsonのAPIが拡張された。顧客のTwitterのつぶやきなども解析し、顧客の性格・特性・好みなどを予測できるようになり、顧客が好むであろう言い回しでコミュニケーションを取れるようになった。

さて、ここからタスクの難易度が増す。

### ○タスク3「デシジョン」

例えば、がん診断支援。症状をもとにどんな治療をすべきかという医師の意思決定、判断の支援をする。幾つか提案する治療法に、90%、45%などといったように「確信度」みたいなものを付けて返す。

現在開発中なわけだが、果たしてWatsonをどの程度賢くしなければならぬのか。そのベンチマークとして人間の医師の判断の正しさを基準にするわけだが、判断の90%以上正しかったという医師はほとんどまれ。通常の医師以上に正しい判断をするというところを目指

す。それでもWatsonは、医師の支援ツールということにはなるのだが。

**——がん診断をした時のロジックは人間が理解できるものになるのでしょうか？**

いえ、それは少し難しいと思う。ロジックは機械学習によって、Watsonの内部で自動生成されるから。ただ根拠は提示するようにしていて、それを見たらどうしてこんな治療方針になったのかは分かるようにはなっている

**——画像の解析はしないのですか？**

現状はできていない。CTの結果は、文章レポートとしてもらっている。数値データもあるし、知見をレポートにしてもらったものを使っている。

Watson開発の次の段階では、画像や、動画も直接扱えるようにしようとしている。医療画像を読影出来るWatsonは研究所で開発を進めているが、現在5年計画の3年目くらいのところ。最終的に読影師の国家試験が受かるレベルまでWatsonを賢くしようとしている。

**——これはマシンラーニング（機械学習）？ Deep Learning？**

マシンラーニングと言っていいと思う。Deepかどうかは定義によるけど、われわれはWatsonをディープQ&Aシステムと呼んでいる。

クリニカル・トライアルマッチングというものにもWatsonは使える。新薬投入前にかかなり長い期間フィールドテストをやり知見を集める。その際にフィールドテストに協力してくれる、症状がマッチする患者を集めるのが大変。患者からすると反対に、自分に合う治療を受けてみたいと思っている。だが、巡り合わない。米メイヨークリニックというところで実施中だが、症状にあった患者を集めることでフィールドテストの期間を短くしようとしている。具体的には、Watsonがテストの条件を読み

込み患者のデータを分析して、患者とテストがマッチする組み合わせ候補を抽出している。

#### ○タスク4「ディスカバリー」

ペーラー医科大学で進行中。論文数7万。P53は、がんの発生を抑制する遺伝子。これが変異を起こすとガンになりやすくなる。P53に影響を与える可能性のある酵素を集めてきている。各々の酵素がどういう特徴を持つかを分析している。全ての酵素を分析することはできないので、酵素のつながりの距離を測るということをしている。文献に登場した言葉と言葉の距離を測る自然言語処理のアルゴリズムがあるのだが、それと同じようなアルゴリズムを使う。つまり、酵素間の距離を数値化、グラフ化する。そうすることで実験をしなければならぬ対象が絞り込める。研究者のひらめきを支援する。

ソフトバンクのPepperをネットワークを通じてWatsonにつなげることによってPepperが知的な受け答えをできるようになる。このためにPepperで動くアプリを開発した。音声認識や発声はPepper自身が行う。Pepperの認識した声はテキスト化されWatsonに送られる。Watson側では4つの異なる仕組みが動く。まずは会話エンジン。テキスト化されたデータで簡単な内容のものは、この会話エンジンがそのまま受け答えする。判断できない内容に関しては、会話エンジンの後ろに控えている3つのエンジンのうちの1つが対応する。3つのエンジンは、まずはクイズに強いエンジンで、これはクイズ番組で優勝したWatson。次に料理に強いエンジン、これはシェフWatson。それとWatson自身について語るエンジン。この3つのうちの1つが対応し情報を返し、それを会話エンジンが会話らしい表現に変えてPepperに返す。Pepperはその表現にしたがって、仕草を合わせ発声するようになっている。

#### ——結局Watsonは幾つ存在するのですか？

プロジェクトの数だけ新しいWatsonが作られえている。知識をミックスした形で学習するのは苦手で、知識領域を限定すれば強さを発揮する。クイズもジャンルごとに異なるWatsonを形成している。

#### ——シナプスというチップを開発していると聞いたのですが、詳しく教えてもらえませんか？

大型コンピュータの一番の問題点は、消費電力が莫大ということ。WatsonはUnixサーバーを並べた大掛かりな仕組みで、ものすごく電力を消費する。なので軽くて、低消費電力、どこでも埋め込めるデバイスを作っていかなければならないと考えている。シナプスはニューロコンピュータの考え方のチップ。プログラミングも今までの考え方と違う。

最初のターゲットは、画像認識、車載カメラなど、1つの目的のデバイス向けに考えている。

チップにはまだできておらず、現状はボードレベルで、色々機械学習させている。

### ——シナプスはDeep Learning？教師なし、教師ありですか？

シナプスは教師あり、教師なしのどちらでもできる仕組み。今は、教師ありの画像認識をやっている。これもある意味Deep Learningなのだが、マシンラーニングとか色んな呼び方をしている。最初それなりの相当数の画像を見せて、これはヒゲがあるとか、目があるとか、人間がタグ付けして教えていく。そうすると、次第に自動的に特徴点の抽出をするようになっていく。そうする中で、男性の写真に対し「これは女性である」と間違った判断をしたら、人間が訂正してあげる。そうすることで次第に正答率を高める、といったことをやっている。

ボストンの爆弾事件の犯人が逮捕されたあと、警察に協力する目的ではなく、実験として犯人の特徴を集め監視カメラの画像を分析させたところ、実際に犯人を正確に特定できたそう。

### ——Watsonが普及すれば、将来的には医者が必要になるのでしょうか？

国立病院機構の医師と、この件に関し、話をしたことがある。彼は、Watsonのようなもので新しい治療法が見つければ素晴らしいし、その結果自分たちの職業なくなったとしても、それが社会のためにいいことなら、それもいいと思うと語っていた。10年後はWatsonを信頼して診断する時代がくるかもしれない。

**——弁護士、裁判官といった職業もなくなるのではないのでしょうか？**

どうなのでしょうね。すべてWatsonで行うのは大変なことだと思う。ただ商取引でコンプライアンスに引っかからないようにWatsonにチェックさせるといった利用のされ方は既に始まっている。

**——Watsonの教師は人間だけなのでしょうか？**

人間ばかりではない。例えば、コールセンターは対応履歴が「教師データ」となる。つねに専門の人間ということではない。

**——Watsonが影響を与えそうな業界は？**

現在、Watsonの開発や導入が進んでいるのは、医療、ライフサイエンス、金融などの領域。金融はかなり進んできている。先ほどの保険の査定の話の他にもいろいろある。IBMとは別にサードパーティがアプリを開発してくれている。Watsonのエコシステムができつつある。コンシューマー向けに健康アドバイスのスマホアプリ、超パーソナライズした旅情報のアプリなどもできてきている。

# 第5章 大阪大学石黒浩教授

ロボット工学石黒教授を湯川塾に招いてお話を聞いた。同教授の発想は大変おもしろかったのだが、残念ながらすべてオフレコを条件にした議論になったので、ここで紹介するわけにはいかない。その代わりに、塾の直前に石黒教授をインタビューさせてもらった取材メモを掲載することにした。

取材日：2014年10月24日

場所：日本IBM（東京・丸の内）

取材対象：大阪大学大学院基礎工学研究科・石黒浩教授

## 石黒教授インタビュー

### ——存在感とは？

Feeling of presence。日本語にしかない言葉。無理やり英語にするとfeeling of presence。そこにいるような感じ。アンドロイドでいうと、アンドロイドに内蔵したスピーカーを通じて僕が喋ってれば、最初は動きがぎこちないと感じても、相手は5分もたたないうちに慣れて、僕だと思って普通にしゃべる。（補足説明：アンドロイドとはこの場合、見た目が人間そっくりに作ったロボットのこと）

### ——見た感じなどを単純化されたロボットでも存在感を感じることができるのでしょうか？

テレノイド（性別もないほど単純化されたロボット）でも高齢者は存在感を感じる。遠隔から音声を送信されていることを分かっている、それには関係なく、そこに人が存在していると認識する。

米朝師匠も本当にそこにいるようにみんな思った。（補足説明：上方落語の重鎮、桂米朝氏そっくりのロボットを作成し、演芸場でそのロボットを操作し落語の音声を流した実験）

つまり、中身が本当の人間かどうかを見る手段がないので、しゃべり方などが人間らしければ、人はアンドロイドを人間として認識するということだと思う。



# 人間に似ていなくても姿を補完する想像力が人にはある

## ——人がロボットを「人間らしい」と認識する最低条件はなんでしょう？

人間の認識は2通り。人の中には知人の〇〇さんを認識するモデルが入っている。しゃべりかた、見た目、などの情報を合わせていき、だいたいそのモデルに合っていれば、不気味と感ずることなく、〇〇さんだと感じて会話を続けることができる。

しかし、その場合に動きがちょっとでも変だと不気味になる。

テレノイドは、姿、形がものすごく省略されているので、形からはだれであるかは認識できない。そうするとモデルの使い方が逆転する。分からないところを人間はポジティブに補完する。（補足説明：テレノイドは遠隔操作するロボットで、人間そっくりに作っているのではなく、姿を簡略化してある）

声から姿、形を想像し、その想像したものを、テレノイドに投影する。

基本的に人間は、相手を想像して対話する。最初だけは相手を見るが、その後は相手をほとんど見ていない。相手がそこにいるものとして会話をする。

## ——人間の想像力が足りない部分を補完するわけですね。そうするとロボットをそっくりに再現する必要もない？

そう。では、どこまで省略できるのかというと、2つのモダリティまで省略できるという仮説がある。

モノを認識するための情報は、形であったり、匂いであったり、重さであったり、いろいろなモダリティがあるが、2つ以上を合わせたときに、そのモノを認識したという感覚になるという仮説だ。つまりそのモノが人であると認識するには、人らしい情報が2つあればいい。

テレノイドは、声は遠隔だけど人間の声、触感は人間ぽく作ってある。その2つの情報が人間らしいというだけで、高齢者はテレノイドを人間と認識する。

ハグビーを使ってもらくと、コレステロール値が低下する。（補足説

明：ハグビーは抱き枕のようなカタチのロボット。顔や身体が簡略化されているので、人間は想像力で補完する。携帯電話を埋め込むことで、ハグビーが話ししている感じになる。高齢者がハグビーを抱いて、遠く離れた家族と会話すると、高齢者のコレステロール値が大幅に下がったという実験結果があるらしい)

携帯電話だけだと、話が終わって電話を切ったあとに、寂しさが残る。ハグビーだと、寂しさは残らない。ストレスホルモンが減って、人の存在を強く感じる。微妙な差ではない。実験では数値が大きく減少した。これで人生が変わったという話は何組も存在する。

最低限のもモダリティとして、声は不可欠。でもにおいも効果がある。

小学1年生に対してハグビーを使った実験でも、成果が出た。幼稚園は保育さんがスキンシップを提供してくれるが、小学校はいきなり集団生活になる。なので不安定になる子供が多い。このためクラスで本の読み聞かせをすると、クラスの前のほうの子供は読み聞かせを聞いているのだが、後ろのほうは全然聞かずに遊び始める。

子供にハグビーを抱かせて、そこから先生の声が出るようにしたら、落ち着きのない子がなくなった。子供たちは、1時間くらい集中できる。ものすごく強く先生の存在感を感じるから。

## 一声がなければ、人間として認識しない？

完璧に状況を作り込むと、声がなくてもだいじょうぶ。

中途半端の状況の作り込みだと、声がないとまったくだめ。

状況次第。ありえない状況を作ったら、みんな無茶苦茶なことを聞く。

ある研究所の受け付けに「いらっしゃいませ」というだけのアンドロイドを置いたら、8割の人間は受け付けの女性がアンドロイドだとは気づかなかった。

受け付けという環境では、質問されることが限定されている。想定の問題に対する受け答えををうまく作り込んでいけば、アンドロイドだとはばれない。

万能の知能ってない。人間は全部状況に依存している。こういう状況ではこう振舞う、こう喋るという知識の集合。どういう場面じゃべら

せるかによって、今の人工知能でも十分に対応できる領域はある。まだ少ないけどね。

高齢者、子供は猜疑心が少なく、認知能力も低いので、今の人工知能を搭載したロボットでも人として認識するケースが多い。

Pepperは普及する。値段が違う。とんでもない値段。パソコンと同じ価格帯。Pepperは、プラットフォームビジネスだ。

パソコンの価格は、ハードとソフトでだいたい半分ずつ構成している。一方でゲーム機は、ハードの価格を低く設定して普及させ、あとはソフトで回収するビジネスモデル。スマホはセンサーも入ってるので、ハードは高いけど、通信料とかで吸収できるモデル。

Pepperもハードを安く設定したモデル。でもパソコンと同じ値段は驚異的。3年契約、保守契約まで含めると50万円ぐらいかかるみたいだが、同じものを大学の研究室で作れば、2000万円から3000万円、へたすると1億円ぐらいかかる。

業界関係者は、みんな昔からロボットのプラットフォームは狙ってた。ただかなり大きなコストがかかるので、だれがいくか、みんな見ていた。Googleはいずれいくとみんな思った。なので孫さんは今しかやるタイミングがないと思ったんじゃないかなあ。Googleが出てくればGoogleの一人勝ちになるのは間違いないので、その前にプラットフォームを取りたいと思ったのかもしれない。

**——「存在感」を感じるまでに5分ぐらいの対話は、不可欠なのですか？**

対話が成立するまで。人間でも顔つきのこわい人、挙動の変な人がいるけど、対話が成立したとたん、挙動や顔つきよりも、対話の内容に引き込まれるので、変に感じなくなる。

**——今の興味は？**

ロボットにはなくて、人。人らしさって何か。心や意識はどうすれば再現できるのか。

**——科学で意識はどの程度わかっているのでしょうか。**

ほとんど分かっていない。医学的には、意識のあるなしはわかっている。覚醒しているかどうか。

哲学的な、高次の意識ってほとんど分かっていない。

## ——ロボットで再現できるものでしょうか？

実は意識や心といったものはないのではないかと思っている。心はどの言語にもあるんだけど、ただだれもクリアに定義したことはない。意識も感情もそう。

人間は複雑なんだけど、まったくでたらめに動いているのではなくて、なんらかの原理、特徴があって動いている。それをわれわれは心と呼んでいるのではないか。お互いに心があると思うけど、実際に心があるかどうか内省できない。でも互いにあると感じる。つまり社会的な相互作用の中に現れる現象としては、心というものはあるが、実際には単なる原理、特徴にしか過ぎないのではないか。

同様に人間のような対話可能な複雑なシステムが、行動の裏になんとか何かぼんやりした根本原理みたいなものを持っていれば、われわれはそれを心と感じるかもしれない。

それを証明するには、ロボットを作って、だれもがこのロボットには心があるって言いだしたら、そのロボットの中身を見ればいいという話になる。何も無いということがわかれば、複雑なシステムで人とコミュニケーションできるものには、ある一定の複雑さを超えると心のように感じるものが表出すると証明できる。

僕の目的はそれ。

## 人間に「心」はない？ あるのは複雑なシステムの行動原理

——心がないということを証明するために、ロボット研究を続けているわけですか。なるほど。でもそのためには人工知能を改良して、より複雑にしていけないといけなのでは？

かもしれない。でも高齢者だったら、すごく単純なロボットにも心はあ

と言う。人工知能を搭載していないロボットなんだけど、心があると高年齢者は言う。

それは対話とは何か、という問題につながる。

対話とは1対1のときは、100%相手の言うことを理解していないと対話が続かない。そう思うじゃないですか。

ロボットと人間の1対1なら、ロボットにかなりの知能が必要。ロボットの受け答えがうまくいかないと、「このロボットは俺の言うことが分からん、バカだ」だと思う。しかしロボットを2体以上用意し、ロボットの輪の中に人間を入れて会話させるは、人間側の対応が変わってくる。ロボット同士で「彼（人間）の言っていることよく分からないね」と話をし始めると、人間側が負い目を感じ始める。対話できないのはロボットではなく、自分ではないのかと思いはじめ。つまり完璧な人工知能でなくても、人間はロボットに対して「心」を持っていると感じるようになるわけだ。

ロボットマジョリティーの世界では、人間は言葉が通じなくてもそんなえらそうな態度は取らなくなるのではないか（笑）。

ロボット研究は、高次のレベルの心理学、認知科学。認知科学や心理学は低いところで止まっているけど、僕たちはロボットがあるおかげで、メタレベルの心理学や認知科学の研究ができる。うちの実験室の連中の半分くらいは最先端の認知科学をやっている。

ロボットがあるから社会関係を再現できる。ないときは内省するしかない。

**——ロボットに対して心を感じるのであれば、ロボットが恋愛対象になったりするのでしょうか？**

女性のアンドロイドの横に座って手を握るとドキドキする。普通の女性の横に座ってもドキドキしないのに、アンドロイドの横に座るとドキドキする。不思議な感覚。研究室の若い連中の中にアンドロイドに対しては、顔を真っ赤にするやつがいて、それを見て、生身の女性のほうが起こるんですよ、ハハハ。

これはアンドロイドとは最初から恋人状態になれるから。アンドロイドにキスをしてアンドロイドは怒らない。人目を気にしなければキスできるけど、人目があるのでしない。これって恋人関係。アンドロイドとは、最初から恋人状態なんです。

恋をすることを、相手を自由にしたいと思うことと解釈するのであれば、アンドロイドはある意味理想的かも。自由にならないのがいい、というのならまた話は違うが。

でも、まったく自由にならないのなら、恋にならない。自由になる、ならないのバランスが、自由になるほうに傾くのが恋をするということだとすれば、人間とのかかわりが苦手な人にとっては、アンドロイドは理想的な恋人になるかもしれない。

## ——ロボットは世界を征服するのでしょうか？

何も大変なことは起こらない。コンピューターがいずれ人間よりも優秀になり大変なことが起こると大騒ぎしている人がいるが、既にコンピューターは人間より優秀。コンピューターはアルゴリズムがあれば、人間より早く計算できる。物事が定義された瞬間に全部コンピューターが勝つ。

「直感」とか「思いやり」とか定義が曖昧なものだけ、人間が勝つと言われるが、実際には勝ってはいない。定義していないだけ。定義がないので、パフォーマンスを計測できない。将棋とか、クイズとか、定義や、プロセス、パフォーマンス計測方法がはっきりしているものは、全部コンピューターが勝つ。ロボットは単にコンピューターの延長なんで、コンピューターで起こっていることがそのままロボットでも起こる。

なので今の世界で征服されていないだったら未来も征服されることはないし、もし既に征服されていると考えているだったら、征服されるだろう。

例えば株の取引は既にコンピューターでやってる。サブプライムの問題もコンピューターが取引してなかったらだいじょうぶだったと言われる。これを、金融業界がコンピューターに征服されている状態と思うかどうか。

僕は共存と思っている。機会と人間は対立構造にあるわけではなく、遺伝子以外に人間の物理的な身体的な制約を克服するための進化だと思っている。

義手、義足を使っている人と戦争することになるのかというと、そんなことにはなりえない。

——では、人間の強みはどのようなものになっていくのでしょうか？

強みはない。なぜないといけない？ そういう概念こそが間違っている。

——では、表現を変えます。これからの社会で人間はどう貢献していくのでしょうか？

人間は存在することにしか意味はない。存在している限りは貢献している。生物学的に言えば。すべての生物の目的は生き残ることなので、生き残っている限りは、貢献していることだし、それが機械の体になろうが、コンピューターになろうが、人間として生きている限りは、すべてが貢献。

——仕事はどうなっていくのでしょうか？

物理的な仕事はどんどんなくなる。仕事や技術開発の意味は人間理解に向かっている。技術は人間の能力を機械に置き換えるもの。どこまで置き換えられるかを試している。それは引き算で人間を理解するようなもの。最後に何が残るのか見ようとしている。技術開発はそういう意味の人間理解だと思っている。

——最後には何が残るのでしょうか？

まだ分からない。単なる情報の流れだけかもしれない。

今あるすべての仕事がどこに向かっているかという、2つしかない。1つは世の中の起源の探求、もう1つは人間とは何かという探求。

原子分子の世界、宇宙の起源、物質の起源。これは特殊な物理学者がやってる仕事。ほかのすべての仕事は、背景に人間がある。取材もそう。僕に興味を持ってもらって取材してもらってるわけだし、工学は人に役に立つものを作るとするのが定義。経済は人のお金にまつわる話。全部、人間に向かっている。人間社会におけるありとあらゆる仕事は、人間理解なしに成り立たない。

人間を理解するために道具を作っている。その道具であるロボットがどんどん使えるようになってくると、道具を作ることよりも人間を理解することの比重がどんどん大きくなっていく。

これからの社会は、本来の人間理解のほうにみな向かうようになる。

将来便利になると、全員が哲学者になると思ってる。

時間ができるとみな何をするのか。本を読む時間もできる。勉強できる時間も増える。

ほかの動物との違いは大きな脳を持っているということ。しかもそれは客観視できる脳。ほかの動物は本能だけに従って瞬間だけで生きているんだけど、人間は違う。

客観視というのは、世の中で自分の立ち位置はどうなっているのかシミュレーションすること。

脳の大部分は、自分をシミュレーションし客観視しようという行為に使われている。

人間の生きている意味はそこにある。

昔の脳はもっと小さかった。脳がどんどん拡張されていく中で、人は人を知るために生きる。客観視できるので、普遍的なものを簡単に作れる。

技術を生む、客観視をする、は同じこと。科学的な視点で物事を見ることができるということは、客観視ができるということ。同時に自分は何であるのかという自問自答にもつながっている。

自問自答するほうが優位で、科学はおまけだと思っている。人間のセンサーはすべて外を向いていて、内側を向いていない。自分の姿形、自分は何であるのかは見えない。

自分を知るために社会的なインタラクションをする。すべての生物は構造的に自分が見えなくなっている。それを見るために社会を構成し、他人を鏡にして自分を理解しようとしている。

その理解の最も強烈な方法が客観視。その結果、生まれたのが科学。

でも世の中、結局主観しかない。科学で説明できるところはたくさんあるんだけど、説明できないところもたくさんある。たとえば心とかは人の主観。どう感じるかは人それぞれの問題。世の中に心という言葉が存在し、同じ概念を共有している。心を説明するには、主観の集合としての客観というものを、説明しないとイケない。今後僕らがやらなければならないのは、そういうこと。



今までは統計的に、主観の集合として数学的に表して非常にリジ  
ットな形で表現していた。これからは、それぞれの主観を尊重しながら  
も社会的な通念を作り出すような、そういう研究が必要になる。

## ——心、意識は進化しているのでしょうか？

ネットの技術とかで情報に対する物理的、空間的制約が取り払われ  
ることで、人間の意識は進化していると思う。昔は地球規模の意  
識なんて持てなかった。

人間の能力が飛躍的に拡張してくると、物理的な身体ってほとんど  
意味をなくすと思う。「攻殻機動隊」はそんなに間違っていないかもし  
れないって思ってる。

人間の定義も、肉体があるかどうかはどうでもいい。義手、義足があ  
っても人間だから。

## ——集合的主観で人間を理解できるようになれば、意識も進化 するのでしょうか？

理解とは関係なく進化する。

正しい理解はないと思う。物理学でも、ニュートンからアインシュタ  
イン、さらには量子力学と、どんどん世界観が変わっている。なので正し  
いものってなくて、それまでの理解はこうだった、ということしかない。

正しい理解をする人はごく一部でいいので、その理解に基づいて新  
しいツールを作る。たとえば量子コンピューターとか。そのことで全体が  
進歩するんだと思う。

正しいことは、多くの人を理解できる範囲を超えている。これから理  
解できる人と、理解できないまま生きていいる人の二極化がさらに進  
むと思う。

情報が増えることで、驚きは減る。でもうつは増えるかも。そのときは  
また新しいイノベーション、人の心をガイドする何かが出てくるかも。

カウンセリングは増えている。肉体労働から解放されて、考える時  
間が増えているから。

そこにイノベーションが起こるかも。脳科学が発達して、脳の活動を

モニターしながら、安定状態に持っていくような技術が近い将来出てくるだろう。ニューロフィードバックの技術は、結構僕ら研究してますけどね。

物理的なタスクをロボットがやってくれると、今度は、精神的なところに技術の焦点が当たるはず。

物理的なことはロボットがやってくれるのなら、人間はロボットがやらないような精神的なところの仕事に移行する。対話サービスとか。

塞ぎがちな状態や軽度の落ち込みなど、昔だったら見過ごされていた心の問題もケアしようということになる。

**——医者が不要になるという話もありますが、どう思いますか？**

医者ではなく看護婦が先に不要になる。医者のほうが力を持っているので。

**——医師の診断よりも、看護師の親切な態度のほうが必要性があるように思うのですが？**

小さい子供にとっては、看護師よりも、ロボットぽいもの、アニメのキャラクターぽいものの方がいい。そのほうが適応しやすい。

認知症患者にとっても、人間の看護師よりも、辛抱強くつきあってくれるロボットのほうがいい。

## 第6章 匿名仮想座談会

取材のあとや、塾の懇親会で、結構いろいろと本音の話が出た。そうした本音の話にこそ価値があったりする。TheWave湯川塾の最大の価値は、本音の議論だと思う。ただ本音は、一般に公開できないものが多い。僕自身も発言者に迷惑をかけてまで公開したいとは思わない。そこで発言者を特定せずに、発言の本質だけを座談会風に創作してみた。これなら発言者に迷惑をかけることもなく、本質的な情報を伝えることができるのではないかと思う。

念のため強調しておきたい。この座談会は完全にフィクションである。それを理解した上でお読みいただきたい。

### DEEP LEARNINGの価値とは

**司会者 Deep Learningについてどう思いますか？ 本当に大変なブレイクスルーだと思いますか？**

**B** 今までできないことができるようになったわけですから、それはすごいことだと思いますよ。

**C** そうですね。顕在意識上でどのようにわれわれが思考しているのかは、もちろん顕在意識上だから分かるのですが、潜在意識下ではよく分からない。多分、何層にも分かれて認識しているんだろうな、という仮説はあるわけですが、その仮説通りに回路を組み立てれば、コンピューターがモノの特徴を自分でつかめるようになった。これはすごいことですね。

**D** そうです。2歳までの赤ちゃんの学習の仕方が分かったわけですが、その部分だけ、今まで分からなかった。なのですごく意味のあることだと思います。

**E** エラー率が大きく下がったというのは確かにすごいけれど、でも大騒ぎするほどすごい話でもないと思う。ヒントン教授が可能にしたのはバック・プロパゲーションというアルゴリズムを分散計算できるようにしたというところだけ。それ以外は何もしていない。従来まったく解けなかった問題が解けるようになったということではない。概念が獲得できたわけじゃない。単にノイズ処理しているだけ。Deep Learningがコンペティションで勝ったのは、3項目のうち1項目だけだし。他の項目は従来のアルゴリズムが勝っている。

**D** まあDeep Learningはまだ新しい技術なんで、そのうちほかの項目も勝てるようになるんじゃないでしょうか。

**司会者** でもこの分野では世界中の注目を集めてますよね。

**E** ほかに何も注目すべきことがないだけ。部分特徴を積み上げて、ちゃんと認識させる。その上に7層ぐらいの階層構造にすれば、パターン認識がうまくいくというのは、もともと分かっていたこと。Deep Learningで初めてできたわけじゃない。

**司会者** Deep Learning自体の意義はさておき、Deep Learningのおかげで世界の人工知能の研究者が、人間の脳を模した回路の研究に再び注目を集め始めたのは事実ですよ。

**B** はい、それは事実です。

**C** その結果、人工知能の研究が急速に進化しているのも事実です。

**司会者** ではこのまま人工知能は一気に進化し続けるのでしょうか？

**B** もちろん課題は幾つもあるでしょうが、このあと大きな壁があるようには見えませんね。

**C** うーん、どうなのでしょう。今はDeep Learningが大きなブレイクスルーとなり、その周辺の研究が一気に進んでいます。2, 3年後にまた新たな壁が出てくるかもしれませんね（笑）。

## 人工知能は人間の脳を再現できるのか

**司会者** 人工知能は本当に人間の脳を再現できるようになるのでしょうか？

**B** 人間の脳は電気信号で情報を伝達しているので、コンピューターで再現可能です。計算はできても人工知能は感情を持たないと一部で考えられていますが、感情も脳

の中で作られているのであれば、感情もコンピューターで再現可能なはずですよ。

**C** 欧米は人間の脳をスキャンして同じものを再現しようとしています。日本の研究者は個別の部位の機能を実現して、それを組み合わせることで脳全体を再現しようとしています。どちらの方法がいいのか分かりませんが、二通りのアプローチで進められています。

**E** わたしは再現は無理じゃないかなって思っている。「飛行機」は作れても「鳥」は作れないんじゃないかな。

**C** いや、少なくとも日本の研究者は「鳥」の再現を目指しているわけではないですよ。われわれは脳の機能を果たすコンピューターの開発を目指しています。「飛行機」の開発を目指しているんです。

**E** 欧米は「鳥」を目指してるでしょ？

**C** そうかも知れませんが。

**E** 欧米の研究は破綻するんじゃないかって思います。だって脳のことはまだほとんど分かっていないんだから。

## これからの仕事について

**司会者** これからどんな仕事有望だと思いますか？

**B** 人間は二極化していくように思いますね。自分らしく生きていく人と、そういう人を使う人の二極化です。仕事はどんどんなくなっていく一方で、生活に必要な物の価格は技術革新を受けてどんどん低くなる。それほどお金を稼がなくても何とか生活ができるようになる。今日でも就職しない若者が増えています。それはそうした未来の兆しだと思いますね。お金がなくても何とか生きて行けますから。ただ仕事はお金を稼ぐためのものだけじゃない。なので儲からなくてもやりがいのある仕事をする人が今後は増えていくでしょうね。

**D** 人間の強みは変化に強いことです。人工知能は、自分で学ぶためには膨大なデータが必要ですから。人間の脳は本当に優秀で、非常に少ないデータでも学習できる。人工知能も今後進化するでしょうが、学習に膨大なデータが必要だということはそう変わらな

いと思う。なので変化に強いという人間の強さを活かした職業は人間に残るでしょうね。

**A** 働く必要がなくなった時、人間は「自分とは何か」を考える方向に向かっていくと思う。そういう意味で働くと思う。自分のやりたいことを通じて、自分を知らうとする。やりたいことだけをやる。そんな人が増えるでしょうね。

**C** うん、全部をロボットにやらせる必要はないものね。やりたいことは人間がやりたい（笑）。

**E** ロボットのお世話を喜んでする人も増えるだろうな。犬型ロボットのAIBOなんかも、お世話すると喜んでくれるようになると、人間は喜んでロボットの世話をするようになると思うね。

## 人間はロボットに征服されるのか

**司会者** 人間はいずれロボットに征服されるのでしょうか？

**C** 人間とロボットを比べること自体が、馬鹿げていると思う。自分に自信がない人たち、ロボットの登場に不安を感じる人たちが、無理やり比べて、それでさらに不安になっている。実に馬鹿げています。

**B** 人間ってなんでしょう。今、それさえも定義が難しい。人間もロボットも定義されていない。義手、義足の人はロボットなんですか？何割まで生身だと人間なんですか？脳と指先しか動かない人は、電動車椅子で移動している。その人はロボットなんですか？自分の頭で考えずに何でもネットで検索して流行に乗ってるだけの人。ネット上の影響力の強い人の意見に左右される人。そんな人って、頭はコンピューターの影響を受けていて、身体が人間。それでロボットより偉いって言えるんですか？人間であるだけで、ロボットより偉いんですか？

## ロボットは自分で学習できるようになるのか

**司会者** 人工知能は、自分で学習し、進化できるようになりつつあるのですが、それでも人間らしい感性を身につけるためには、人間と同じような経験を積む必要がありますよね。そのためにはロボットという「身体」を持たせて、いろいろ体験させなければならぬと思いますが、現時点でロボットは自分で学習できるものなのでしょうか？

**C** ロボットが自動学習して自分で微調整する機能はまだないですね。やはりティーチングしかないと思います。人工知能を搭載する前に、人間が思った作業をロボットにすぐに

伝えるということが大事。まだ人工知能を搭載する段階ではないように思います。

**B** 人間の感情って、相当の期間の経験の積み重ねで確立してきたものだと思うんです。明確な理由もなく、「きれい」「怖い」とか感じるようになるには、何世代もの経験の蓄積が必要なんだと思うんです。これをロボットで再現するのは非常に難しい。

**司会者** じゃあ、本能は人間がロボットに埋め込むしかないんでしょうか？

**B** まあ埋め込んだとしても、大雑把にしか埋め込めないのが難しいと思います。

**司会者** では人工知能にバーチャルリアリティの中でいろんな体験をいっぱい積ませることによって、感情、本能というものを獲得させることは不可能でしょうか？

**D** ありふれたバーチャルリアリティではダメでしょうね。ものすごくリアルなバーチャルリアリティがあれば別かもしれないですが。

**F** リアル以上にリアルなバーチャルリアリティが欲しいですね。下手なリアルな経験よりもバーチャルにいた方が感覚豊かになる、というようなやつ。

**D** それってもはや「バーチャル」と呼べないかもしれないですね。リアル以上にリアルで、リアルに生きるより学びの速度が速い空間。

**F** バーチャルリアリティも最後はそこに行き着くような気がしますね。

## ロボットが奪う仕事、奪わない仕事

**司会者** 工場以外のところで、どのような領域にロボットはこれから入っていくのでしょうか？

**D** 廃棄物処理作業などには、すぐにでもロボットは活用されるでしょうね。

**B** バイオテクノロジーの研究所などでもニーズがあるって聞きます。試薬を試験管に一滴落とすだけのために日曜日に研究者が出勤しなければならないケースとかがあるそうですから。

**司会者 飲食業とかには、まだロボットは入っていかないんでしょうか？**

**F** 牛丼屋だけ、というように限定すれば可能でしょうね。ただ作業工程をロボットに最適化したものにしないといけません。店内もロボットに最適化する必要がある。牛丼屋チェーンがそうした投資にどれだけ積極的か、またお客さんがそれを受け入れるかどうか、にかかっていると思いますね。

**司会者 介護ロボットとかはどうでしょう。**

**E** まだしばらくは難しいかも。ロボットの動きで人を怪我させる可能性がまだあるから。

**D** ロボット導入を前提とした作業工程を考えていく必要があるでしょうね。あと、ユーザーがロボットに簡単に教えられるようなインターフェースが絶対に必要。介護の現場の人たちでプログラミングできる人は非常に少ないから。会話ができるようにするために人工知能を搭載しないとイケないでしょうね。

**A** 将来的には、介護はロボットの仕事になる。だって人間はやりたくないから。特に高齢者の話し相手には、ロボットのほうが辛抱強くていいと思うよ。

**司会者 ソフトバンクのPepperはどう評価しますか？**

**D** 僕はまだ購入しないね。初期不良があるかもしれないから。

**B** いや、発売になるころにはだいじょうぶになるんじゃないかなあ。それに業種によっては早い時期に導入したほうがメリットが大きい業種もあるし。

**D** でも、20万円で発売するというのがすごいよね。解体して部品を売っても、20以上稼げるんじゃないかな（笑）。

**B** あと何年かすればコストが下がり20万円でも見合う時代になるでしょうが、そうなればベンチャー企業を含め競合他社が一気に家庭向けロボット事業に参入してくる。そう



なる前に大量の資金を投入して参入する。確かにベンチャー企業と差をつけるには、その方法がいいでしょうね。

**E** お掃除ロボット「ルンバ」の開発者に話を聞いたんだけど、ルンバを発売しようとしたら顧問弁護士から「ユーザーがつまづいてケガをすれば訴えられる。やめたほうがいい」とアドバイスを受けたい。そんなこと言ったらロボット事業なんてできないって、アドバイスを無視して発売したら大ヒットになった。メーカー、特に大手になると、そうした訴訟リスクを考えて身動きできなくなりますよね。その点、ソフトバンクはやっぱりすごいなあ。他の大手メーカーでも技術的に作れないことはない。でも、みんな訴訟が怖くて動けない。

**B** 家庭向けロボットはGoogleが先行するって思っていたのに、ソフトバンクが世界に先駆けて発売しようとしている。すごいよね。Googleはロボットベンチャーを数社買収しているけど、ロボット事業担当の重役が突然、退社している。どうやらベンチャーのCEOはみな天才肌で、そうした数人の天才をまとめるのが大変だったのではないかって噂になっている。船頭多くして船山に登るということかもね。その点、ソフトバンクは一人の天才が指令を出すので動きが速い。

**C** そうだね。100年後には一家に一台ロボットが普及するって2010年に講演した次の年には、もう動き出していたみたいだもんね。

**F** 日本が世界に先駆けて「ロボット 1 家に 1 台」時代になれば、うれしいよね。Pepperのおかげでロボットに取り組み日本人技術者は間違いなく増える。IBMの人工知能Watsonとも連携するみたいだし、日本にとっては素晴らしいことだと思う。ソフトバンクにはがんばってもらいたい。

**D** ただソフトバンクにとっては赤字垂れ流しの事業になるわけだし、ソフトバンクが体力的に持ちこたえられるのかが課題だね。発売 2 年ぐらいで、それなりの成果を出せるのかどうか。Pepperが失敗したら、日本から次に続こうという動きは出てこないだろうし、「ロボット大国日本」の夢も遠のくだろうなあ。

**E** 日本はロボット事業を展開しやすい土壤があると思いますね。ロボットを友達だと認識している国民性がありますから。また初音ミクのように、コンテンツをみんなによってたかって作っていくという文化もある。

**司会者** Pepperの機能でどんなものを期待しますか？どんな付加価値サービスなら月額使用料をさらに支払ってでも受けてみたいですか？

**D** テレビのリモコン代わりになってくれたり、玄関のインターコム映像を写しだしてくれたりしたらうれしいかも。

**E** テレビのチャンネルを変えてくれるより、リモコンを探して持ってきてくれたらうれしい。

**B** 老人向けコミュニケーションサービス。月額2000円くらいまでなら高齢者の相手に特化した対話能力に対して支払ってもいいかも。それが可能なら離れて一人暮らしする母親にPepperを一台プレゼントしたい。

**C** でもロボットって特に役に立たなくてもいいと思う。「弱いロボット」でいいんじゃないかな。Pepperのローンチイベントで、[Pepperが途中で歌詞を間違ってた](#)のがあったけど、あれ見てPepperに愛着がわいた。コミュニケーションロボットの本質って弱いロボットだと思ったよ。

## 司会者 意識って何なんでしょう？

**A** 医者が言うところの「意識がある」「死亡した」というような意味なら、「生きていること」という意味でいいと思います。でも哲学的な意味での「意識」ってなんなのか、だれにも分からないと思いますね。

**D** 受動意識仮説という仮説があります。実は、われわれの多くの動きって脳で意識するよりも若干早く身体が先に動いているといわれます。脳に電極を貼り付け、時計の秒針を見ながら指を動かす実験があります。指を動かすと脳に走る電流が見れるようにした実験です。秒針が9の位置で指を動かそうと思ったとすると、脳には6の位置で電流がはまっている。自分が意識の中で指を動かそうと決める前に電気信号は出ているわけです。でも、自分では自分で動かそうと決めたと記憶しています。つまり意識って単なる条件反射の後付の記憶ではないかと思うんです。ではなぜ意識、後付の理由づけが必要なのかというと、それがあると少量のデータで学習できるわけです。禅では「沈黙で対話せよ」と言いますが、禅を受動意識仮説的にいいかえると、条件反射、無意識が主役。意識は学習効果をはやめるための副産物。だから意識にとらわれるなということなんだと思います。学習を速くするのが意識の目的なんだから、それ以上にいろいろ思い悩むなっていうことだと思います。

**B** 確かに脳の大半が条件反射で動いています。自分が決めた意志を後付するが受動意識仮説です。内的体験は、受動意識仮説でいうところの自分の意志です。意志は意味付けすることによってストーリーができる。ストーリーが出来ることによって記憶が出来る。記憶が出来るようになって学習が早くなる。そんな風になってるんじゃないかなあ。トラウ

マは機械学習であればおきない。なぜならデータが少な過ぎて学習できないから。人間は少ないデータ、1回の経験でも学習ができて、反対に過学習したのがトラウマということなんだと思います。ロボットに意志を持たせることで学習を早くしようという動きはまだないですね。これからそういう研究も始まるかもしれないけど。

### 司会者 「直感」ってどこから来るんでしょう？

**C** 直感って、パターン認識しておいて、色々な情報を瞬間的に組み合わせる能力のことだと思う。たくさんのパターンを取り込んでおいて、いきなり法則を見つけたり、瞬時的に取り出す能力を、直感って言うんじゃないかな。

### 司会者 「直感」は人工知能でもできちゃう？

**F** ある程度できるでしょうね。人間とまったく同じ脳の機能持てば、人間と同じようにできるかもしれない。しかし！人間の中で「直感」を持って世の中を変えている人はごく一部。トップ中のトップだけ。

**C** でも直感にはいろんなトリガーがある。身体経験や過去の経験などから直感が沸くことがある。脳だけでは直感にはわからないんじゃないか。

### 司会者 シングularity後の社会はどのようになるんでしょうか？

**A** 生身の脳では理解できないほど変化するというのが、シングularityの定義だったりするので、今われわれには想定できないんじゃないのか（笑）。

**B** エネルギーも太陽光発電で十分に取れるようになれば、すべてのものが無料になり、だれもが3Dプリンターを持って、好きなものを好きなときに無料で作り出せるようになる。そうなると貨幣経済は終るよね。

**C** すごく生産効率がよくなるし、人が死ななくなるとも言われている。その時、人は何を欲しがるんだろう？

**F** みんな今の無職の若者と同じような感じになるんじゃないか。「生きていけるから頑張る必要ない」みたい。

**D** そうなったら、人間は何を求めていくんだろう？

**A** 知識しかないと思う。人や自分についての理解を求めるんだと思う。「なんで自分は生きるのだろう」と悩み出すんじゃないか。

**B** 「人間」の定義なしにシングularityを語っても意味がない。愚の骨頂。

**C** でもシンギュラリティが「可能か?」というと、可能だと思う。

# おわりに

ちょっとばたばたの状態で書き終えた。誤字脱字や事実誤認は、取材先や仲間にチェックしてもらったが、完璧ではないかもしれない。電子書籍なので、指摘があるたびに修正していきたいと思う。

この電子書籍の作成に協力していただいたみなさんにお礼を申し上げたい。特にこのプロジェクトの先頭に立って仕切ってくれた鈴木まなみさん、校正に力を発揮してくれた谷口正樹さん、電子書籍のマーケティングに尽力してくれた茨木友幸さん、講義メモを作ってくれた奥山浩通さんに感謝の言葉を述べたい。

2015年 1月21日

湯川鶴章

## 著者プロフィール

ITジャーナリスト。学習コミュニティ[TheWave](#)代表、TheWave湯川塾・塾長。株式会社あしたラボラトリー・チーフストラテジスト 1958年和歌山県生まれ。大阪の高校を卒業後、渡米。米カリフォルニア州立大学サンフランシスコ校経済学部卒業。サンフランシスコの地元紙記者を経て、時事通信社米国法人に入社。シリコンバレーの黎明期から米国のハイテク産業を中心に取材を続ける。通算20年間の米国生活を終え2000年5月に帰国。時事通信編集委員を経て2010年独立、ブログメディアTechWaveを創業。2013年から、編集長を降り、新しい領域に挑戦中。著書に「未来予測：ITの次に見える未来、価値観の激変と直感への回帰」（2013年）、「次世代マーケティングプラットフォーム」（2008年）、「爆発するソーシャルメディア」（2007年）、「ウェブを進化させる人たち」（2007年）、「ブログがジャーナリズムを変える」（2006年）。共著に「次世代広告テクノロジー」（2007年）、「ネットは新聞を殺すのか」（2003年）、「サイバージャーナリズム論」（2007年）、「閉塞感のある君へ。こっちへおいでよ。」（2013年）などがある。

より詳しい情報は以下のメディアでどうぞ。

BLOGOSメルマガ「[ITの次に見える未来](#)」

有料オンラインサロン「[湯川鶴章のオンラインサロン](#)」

ブログ「[TheWave](#)」