

(2016年6月22日講演)

AI×データはビジネスをどう変えるか？

ヤフー株式会社

チーフストラテジーオフィサー 安宅和人委員

経産省の審議会等で話している内容をお持ちした。今起こっている変革はどういうことで、結果、どのような未来が来ると私が認識しているか、それから、国として何をやるべきかと思っているかという話をできればと思う。

私はもともと脳神経系の研究をやっていて、脳の遺伝子等を見ていたが、ひょんなことから20年ほど前、当時、サイエンティストの多くがほとんど知らなかったマッキンゼーに入社した。そこでももとの関心領域である人間のパーセプションのテクノロジーと言うべきマーケティングエリアに出会い、新商品開発やブランドの立て直しなどに従事していた。もともと科学者になりたかったこともあり大学に戻ったが、恩師がもう退官していたこともあり、アメリカの大学に行ってPh.D.を取った。順風満帆のつもりだったが、ポストドクをやっていたら9.11のテロがあり心が折れてきて、古巣から悪魔の誘いのようなものがまたあり(笑)、戻った。8年ほど前にヤフーから声掛けがあって移った。最初は役員直下で社長室長のような仕事をやっていたが、4年前に体制が変わり、今CSOをやっている。世の中全般をどう見るか、会社としての戦略的リソース配分、幹部の相談役のようなことが主たる仕事である。データ利活用系、基礎R&Dを1、2年前まで統括していて、今も深く関わっている。その絡みで、データサイエンス協会を丸山先生らと一緒に立ち上げた。現在、私は理事かつスキル委員長もやっており、応用統計学会の理事でもある。

その延長で、人工知能やビッグデータとは何かなど、啓蒙的な活動にも従事している。経産省の産業構造審議会 新産業構造部会、文科省の高度データ人材のあり方のような公的な検討にも関与している。文科省系は丸山委員と今年の今ごろ委員として入っていたが、人工知能技術戦略会議の年間100億円、10年間というプロジェクトも、実はここでの検討がベースの1つになっている。また、OECDの50周年のパネリスト、TEDxなどでも話しているという人間である。

新しい国富の方程式 (資料 p.2~14)

この10年、20年で劇的に変わったことは、コンピュータを持ち歩くようになったことだと思う。1985年当時世界で一番速かったCRAYのコンピュータの20倍以上速いコンピュータが皆さんのポケットに入っている。iPhone6・6sなどである。とんでもない時代で、当時のスパコンより速いコンピュータがポケットに入っているという時代である。

併せてネットにつながる機械が爆発的に増えて、なおかつブロードバンド化が劇的に進んでいる。過去 10 年あまりで 1,000 倍以上。これは例えば新幹線が 1,000 倍速になると、月に 2 時間で行ける（笑）。人類史上誰も体験したことのないような変化の中で我々は生きていることはほぼ確実である。結果、片っ端からログ情報が上に上がって行って、いわゆるビッグデータが無限に出てくるという時代である。今もデータはほとんど垂れ流しで、これからも垂れ流し度は高まる一方であるが、コンピューティングパワーは皆さんご案内のとおり激増している。ムーアの法則は以前ほど注目されていないとはいえ、チップ当たりの処理スピードは 10 年で 100 倍ぐらいにはなるし、チップの数も激増している。しかも、大量データ処理技術はコンピューティングパワーのおかげもあり、劇的に上がっている。コンピューティングパワーのためにこれまではほとんど計算的に無理だった Deep learning のようなものも動くようになったことはご存じの通りである。大量データ可視化の技術も急速に進化している。

このようなことを重ね合わせると、産業革命のときに人間と家畜が肉体労働、手作業から解放されたようなレベルの変化が今起きているのではないかというのが、基本的な私の認識である。みなさまも同様のご認識だと思う。退屈な数字入力やハンドリング、Excel ワークなどの簡単な仕事はかなり消え去るだろうということで、歴史的な局面であると考えている。

1 人当たりの生産性を万年単位で見たバークレーの研究者がいる。彼によると、ローマ時代から産業革命まで 2000 年掛けて 2 倍ぐらいしか生産性が上がっていないが、そこからの 150 年ぐらいで 50 倍、100 倍に上がっている。ただ、よく見ると近年、若干成長が伸び悩んでいる。これがもう一発跳ねるのではないだろうか。なぜなら、産業革命のときに労働から解放されて跳ねたのであれば、今の我々の労働の中心である情報処理系が解放されたればもう一発跳ねてもおかしくないからである。

ここからすべての産業が ICT 化していくのはほぼ確実である。世界のマーケットキャップのランキングで 2 月 2 日に非常に歴史的な瞬間があり、トップ 4 がすべて ICT 企業になった。5 番目のバークシャー・ハサウェイ、バフェットの会社も皆さんご存じの通り、実質的にデータをやりとりしている、ある種の ICT 企業といえる。このような会社が上に来て、エクソンモビールのようなかつての産業革命の覇者でも 6 位くらいである。世界最大級のメーカーであるトヨタの前後に Amazon やアリババがいるというような実態である。実際トヨタは 20 位以下で、これだけ大きなものづくりをしても、トップ 20 にすら入れないという時代が今やってきている。

さらに、付加価値の総和である GDP を見ても、過去 20 年ぐらい ICT セクターの中で、実は日本は成長していない。つまり、マーケットキャップの話だけではなく、GDP のような世界で見ても、やはり ICT が富を生み出す中心になっていることはほぼ確実である。しかも、特徴的なのは、利益に対するマーケットキャップを見ると明らかに ICT プレーヤーは高い。これは「未来を変えている感」があるかどうかマーケットキャップに直結している。

以上をまとめると、国富を生む方程式が変わったというのが基本的な認識である。何となく物を造って付加価値を大量に生み出すこれまでのやり方では、富は生まれにくいような仕組みに今なりつつある。

ビッグデータの本質ともたらず変化（資料 p.16~29）

ここで2つの大きな技術的な要因、データとAI的な話、それぞれについて私が思っている重要なポイントを話したい。

マーケティングの世界だと、どのように選ばれているのかを理解するために、ユーザーを見て、ユーザーから聞いて、そしてまたユーザーのつもりで考えるということを行う。この視点で考えたときにどういう手法が通常あるかというところ、「見る」「聞く」を定性的にやっているのがいわゆる定性調査、それをベースに「聞く」を定量化するのが定量調査で、この定性調査で「見る」ことに集中しているのがいわゆる行動観察。マーケティング的な視点で見たビッグデータは基本的には定量的に「見る」方法と位置付けることができる。

自分のようなストラテジストは、どういう利用が、どういう文脈で、どういう意識を持った、どういう人たちにとって行われているかという情報を必要としている。従来型の調査はすべて一気に取るというタイプの調査だったわけであるが、どれほど大きくても5万、10万で、日本人の1000人に1人、1万人に1人ぐらいの部分データでしかなかった。ただ、ビッグデータは基本的にはデバイスを持っている人の全オケージョンを取るのだから、データのカバレッジが桁違いに大きいというところがある。これが大きな特徴だとデータを扱う立場として感じている。

例えばヤフーで持っているあるアプリの緯度経度情報をプロットすると資料 p.19 のようになる。これは地図にプロットしているのではない、緯度経度を打つと、限りなく地図に近い像が出てくる。確かに地図そのものとは若干違う。瀬戸内海がないとか、東京湾も半分ぐらい陸地であるとか、北海道は半分海だったとか。しかも、ここを船が通っていたとか、この辺りにイカ釣り漁船がいっぱいいるという、サンプルベースの情報では見ることが困難なパターンが見える。このロングテールが見えてくるというのが特徴の1つである。テールの長さは大半の人に理解されていない。例えばヤフーの検索は大体 PC で 60%、スマホで 40% 近く、全検索でも 50% 強持っているが、毎日コンスタントに使われるクエリ（検索ワード）は大体 1 日約 60 万種類である。年に 1 回でも使われるとなると、80 億種類もある。つまり、テールというのはヘッドより、少なくとも 4 桁多く存在し、3 つ先のビルまで伸びるほど、すさまじく長い。ここに今まで見えていなかったパターンが埋め込まれている。これを利活用するゲームがビッグデータ活用の一つの特徴である。

もう一つの特徴は、データを入手してから使えるタイミングである。これまでのデータは、入手してクリーニングして、使えるのはせいぜい翌日以降だったが、ビッグデータは基本的には即座に使う。我々にしても、Google にしても、広告は大体 100 ミリ秒ぐらいで返している。100 ミリ秒の間にマッチングを行って返す。非常にフレッシュな利活用をしているわけである。実験的なデータを見ていると、この返すスピードが 10 倍、100 倍遅くな

ただだけで急速にクリック率のようなものが落ちる。言い換えると、フレッシュネスが富を生み出している。このレスポンスのスピードが10分の1遅くなっただけでGoogleの株価は多分10分の1になる。そのぐらいにこのフレッシュネスが大事で、ビッグデータにおけるフレッシュネスの重要性は、強調し切れない。すさまじい勢いで鮮度が落ちる。

ビッグデータの話はデータの大きさばかりに目が行ってしまうが、実は全量性のためにテールのパターンが見えるところ、リアルタイムに使えるところに本質があると考えている。

データ利活用の流れは、脳もコンピュータも同じで入力、処理して、出力する。「脳というのはある種のルーターに過ぎない」。これは先日、日立の技師長の矢野氏と対談したときに言われて名言だなと思ったが、まさにそうである。つまり、一体何につなげるのかということの意味を持ってくる。どのようなアウトプットにつなげるかをしっかり考えないと価値が生まれない。

実際よく考えている会社がある。昨年Googleがおよそ\$60M投資したKreditechという会社があるが、この会社はいわゆる街金融で位置情報やeコマース情報を使って与信し、金を貸している。位置情報を見れば、この人がきちんと定職に就いているか、歌舞伎町で変な店に入り浸っていないかなど、大体分かってしまうので、それで与信の可能性が非常に高まるわけである。なおかつeBayやAmazonで片っ端からログインさせるので、その人がきちんと払っているかどうかも分かる。本当に今までとは全く違う、クレジットビューローなどのデータを使わない与信を掛けている。

今までの消費者ローンはそもそも7割方の人は受けられなかったが、全員を審査の対象にでき、その与信のスピードたるや驚くべきもので、数日掛かっていたものが数十秒で返ってくる。日本でも無人の審査機械があると言われるかもしれないが、実は後ろに人がいて審査している。このKreditechは本当に無人なわけである。ビッグデータで回している。莫大な機械学習を走らせている。

他にもRSMetricsという非常に面白いアナリストレポートの会社がある。これまでは企業の決算報告をベースにアナリストが一、二週間から数カ月を掛けてレポートを出していたわけであるが、この人たちは超高精細の衛星写真をリアルタイムに引き上げてきて、例えばそのWal-Martに客が入っているかいないか、工場にトラックが多く出入りしているかなどを見る。そして、本当にリアルタイム情報をベースに、あのチェーン店はいい、悪いというのをローデータで上げていく。このようにメッシュがこれまでとは桁違いに細かい情報をベースに、とんでもないデータに基づくサービスが始まっている。これも典型的なビッグデータ利用である。

このように、桁違いにメッシュが細かくて新鮮な情報が出てきて、使えるようになったというのが我々の世界を変えているポイントで、世間の話題はすぐAIの話にばかりいってしまうが、実は私は今世の中を変えている最も本質的なポイントの一つはデータだと思っている。

AIの実体ともたらず変化（資料 p.31～75）

では、AIのほうは大事ではないのかというと、大事である。このテーマそのもので、昨年非常に歴史的なサービスが始まった。Google photos である。これは Deep Learning を実装してラージスケールで始まった世界で最初のサービスである。千万単位のユーザーがいきなり使い始めた。上げておけば、勝手にアルバムを作ったり、コラージュを作ったり、静止画をつなぎ合わせて動画にするようなことを機械がやってくれる。このスピードたるや驚くべきもので、手計算ではあるが、人間より 2 百万倍程度速い。200 万人雇っている会社はどこにもないわけである。巨大な日立グループやトヨタグループでも数十万の規模であるから、もう無理である。要は人間には絶対不可能な、どのような会社でもできないことを彼らは機械にやらせているということで、とんでもないことである。

このように、人間がやってきた情報処理を機械とソフトウェアで回しているのが AI である。ただ、これについて非常に危ない誤解が社会を覆っていることに危惧を抱いている。Deep Learning は多用途に使える万能 AI という考えだ。これは革新的ではあるが単なる基礎要素技術の一つに過ぎないわけで、実態としては、非常に高いコンピューティングパワーを持つ環境に、必要な情報科学を実装して、膨大なデータでトレーニングすると特定の AI になるということに過ぎない。これを実現するための技術の組み合わせたるや、我々の音声入力 of 背後の仕組みなどを見ると、とてつもなく複雑なものである。だが、基本的にはコンピューティングパワーと情報科学とデータの三位一体の議論だ。これをとにかくリマインドし続けるのが、この一、二年の私の仕事になっている。実際、情報処理側でも非常に今激しい進化が起きている。Deep Learning のような非常にヘビーなエミュレーションをやっているようなものを、ハードで置き換えるためのチップ、NPU(neural processing unit) が世に出る寸前であり、変わりつつある。この計算力、アルゴ、データの三位一体の変化であるということを読み間違えると大変なことになるなと思う。ということで、ビッグデータ、機械学習は相互に入れ子の構造である。

私が認識している機械学習ベースの AI がもたらす自動化には大きく 3 つある。一つは「識別」である。もう一つは「予測」で、3 つ目は「実行」と言うべきエクスキュージョンに近いタイプのものである。学習ベースの実行ともいえる。

「識別」は先ほどの Google photos の写真の仕分けのようなものが典型的で、Facebook で勝手に顔のタグなどが付いてくるのと全く同じタイプの識別になる。ヤフーだと膨大な量のニュースが流れ込んできて、記事ごとに写真をクリッピングしていかないとイケないが、人間がやるとキリがない。したがって、Deep Learning を使って切り取る場所を自動判断している。これはまさに識別能そのものである。Mirriad というドイツの会社は、動画に入っているような物体を認識して差し替える技術を展開している。彼らはこれは Deep Learning ベースだと言っているが、本当かどうかは分からない。人間がやっているにしてはうま過ぎる。例えば非常に高級な車をアフォーダブルなクルマに置き換えたりするわけである。Amazon も AI で問題のあるレビューは全部外していつている。当社でもヤフオクにどうしても悪い人がやってくるので、そういう人は機械学習に掛けて外していくことを

リアルタイムでやっている。

また、歩容解析 **gait analysis** という技術があり、コンピュータビジョンの一種であるが、人間の歩き方を三次元でレコーディングすると、実は指紋のように決まっている。顔などを見なくても、その人だとほとんど分かってしまう。更にこれを使えば、その人が認知症であるとか、危険な人間であるということなどが分かってしまう。200人ぐらい歩いているなかでこの人が87%の可能性でホシであると分かる。そのため、この技術はまさに今、公安調査庁や、アメリカのNSA（国家安全保障局）などが日々の業務に活用していると言われている。これは危険な状態である。顔を隠していても、このお店に入っていたと分かるという恐ろしい時代がやってきて、悪いことはできなくなる。

さらに、医療では生検などのスライドを見ての診断精度は、ある種のガンでは昨年専門医を超えてしまった。どこかのセミナーで見たが、アメリカの **Massachusetts General Hospital** で皮膚科の実験を行ったところ、MGHの世界レベルの医者の診断の正誤率を超えてしまったというデータを見た。我々が気付かない病変のパターンなどを見いだしてしまう。ゲノムの解析などはもともと人間ではできないというか、我々両親から2ゲノムもらっていて、1ハプロイド当たり30億塩基対もあるわけで、人手ではなかなか解析できない。ものすごい大きなデータベースで解析しないと動かない、そういったものというのは、まさに機械で解析すべき対象であり、これはAIの出番に変わっていくところだと思う。

「予測」の数値予測については需要予測や与信のスコアリングのようなものは普通に行われているが、選挙結果の予測等は、実はAIを使わなくても、私のところで参議院選を予測したらいきなり96%の確率で当たってしまった。それがデータであるが、それにさらにAI的なモデルを突っ込めば、もっと当たるというだけの話である。

今ホットだなと私が思っているのは、心の予測である。気持ち予測が今一気に進んでいる感がある。これを使って広告のマッチングなどがどんどん行われている。例えば **Google Glass** のコンセプトビデオをご覧になった人もおられると思う。これが熱いのは、コンテキストを読んでいることである。ある地下鉄の駅に人が向かうと、その人が何も聞いていないのに、この地下鉄は閉まっていると教えてくれる。ならば歩くということで、意図を検出してウォーキングルートを出すことをやるといった研究が進んでいる。

2年前に **Amazon** がとんでもないパテントを取った。2014年6月18日のニュースであるが、これは何かというと、客が発注する前に出荷するという訳が分からないパテントで（笑）、このパテントが出たとき、我々の業界はかなり騒然とした。本気なのか **Amazon** と、そこまでできるのかと。しかし、今年の夏ぐらいからこれが本当であることが歴然としてきた。例えば昨年、私がした **Amazon** への発注は自動で発注されて、今まで5分以内だったらキャンセルできたが、10秒後ぐらいにキャンセルしようとしてもすでに出荷されたということではできなかった。あり得ない。だが、彼らも嘘をつけない。少なくとも倉庫間の横持ちぐらいしていなければ言えないわけで、実際彼らが予測して物を動かし始めていることは間違いない。まさにAIによる予測の実例である。

マウスやスマホなどの指の動きを見て、物を買うかどうか予測するサービスをエモーシ

ジョンインテリジェンスという会社がやっている。10秒のホットスポットでその人が関心のあるクーポンを出す。コンバージョンがいきなり3倍ぐらいにバンと跳ね上がる。クーポンをそのホットスポットのタイミングでのみ打つので、ベタ打ちと違って利益がはね上がっていく。この会社は年率4、5倍のスピードで伸びている。1年前で月間数百億のコーマースに関わっていた。今は月千億ぐらいにまで行っているのではないかと思う。

こういったことを使いながら Facebook の M や、Google の Now on Tap のような PA (Personal Assistant) がどんどん出てくる。これはまさに予測の話である。

自動化の3つ目の「実行」では、非常に注目を集めているが、人間しかやらなかった翻訳や作曲、デザインなどの自動化サービスがどんどん出てきている。例えばアメリカの Tailor というロゴ作成サービスでは、どのような感じのサービスか、どういう名前にするか、イニシャルベースにするか、どういうスタイルが良いかなどを入力すると、10秒ぐらいで完成する。どこかオリンピックのロゴで非常にもめていた国があったが(笑)、デザイナーはこういうものを使って工夫する時代になるのではないかと思う。デザイナーが「こうではなくて、こう」、「わかる?こうなって、こう」とやるのが人間の仕事になってくるのではないか。

Google DeepMind の DQN は非常に面白い、深層強化学習というスカラー量の目的関数を与えて、ひたすら学習させておくと、多くのテレビゲームで数時間で人間のプロを超えてしまうというとてもない代物である。開発者も気付かなかったような答えを見つけるというので驚いていたわけであるが、世界が一番驚いたのは魔王とまで言われた碁のイ・セドルがほとんど完敗に近い形で負けてしまったことである。数カ月前、日本の井山六冠にまで勝ってしまうような人である。本当に人間技ではないほど強い人が負けてしまうというところまで来ているのが、まさに典型的な今の実行フェーズでの変化だと思っている。

さらに、人間しかやれなかったことの一つにピッキングがある。Amazon がピッキングロボット大会をやっているくらい、ものすごい勢いで自動化が進んでいる。情報処理的な業務はもう明らかに桁違いに向上することは確実で、その上でどうやって付加価値を生むのかが我々の勝負ではないかと感じている。

シゴトの未来についての議論も頭の痛い話で、年中聞かれる。仕事の半分ぐらいは危ないという、Oxford の3年前の論文のせいで質問が多い。そもそも30年続くような仕事など過去にあったのかという基本的な知見に基づけば、そのような話は当たり前であるが、何か大騒ぎになっている。しかも、Google の Larry Page は、十中八九の人は今の仕事を続けようと思わないだろうとか、Tesla/SpaceX の Elon Musk は、AI は人間にとって最大の脅威であるなどと言ってしまって、AI の侵略?のように聞こえて、少しどうかと思う。

機械学習ベースの AI についてあまり語られていない不都合な真実について触れたい。そもそも、AI には少なくとも我々人類が持っているような意志がない。この山に登る、登らないとか、こういう感じの会社や事業にしたいということを考える力は今のところないわけで、これが人との根本的な違いであるというか、幻想であると思う。

人間のように知覚できない。これも言うまでもないわけであるが、我々のようなフィジ

カルな存在がないから、物に触れた感覚のような知覚ができない。色も、多くの人が誤解しているが、物理量ではない。3 ピークを持っている色覚ピグメントが人間の場合たまたま目の中であって、そこから入ってきた入力混合されて、脳の中で知覚として統合されて、我々には色が見える。色というのは脳の中にしかない。味もそうであるし、触感もそうであるし、実はこの身体を持って知覚するというのが基本的な我々の大事な仕事である。ちなみに人間の色覚の原色はたまたま 3 つで、犬は大体 2 つだと言われている。ハトに至っては 4 つか 5 つあると言われている、ペンタクロマティックというとんでもない世界で生きているのであるが、我々の脳では理解できないことも同じ話である。

事例が少ないと対応できないのは、先ほど説明したような仕組みであるから当然である。我々なら体を守るべくへびを振り払うような本能的な動きがあるが、AI には生命的本能は当然ない。

コンテキストを読むのは今のところ非常に不得手である。意味が理解できない。

ヒラメキがない。紙を 4 つ立てれば箱になる、丸めれば望遠鏡の遊びができるなど、人間だったら教えなくても気付くようなことに気付かない。当たり前なことに気付かない。

AI はいわゆる常識的な判断も得意ではない。同じ部屋でミーティングをしていて人が一人二人入れ替わるだけでコンテキストが変わっても、我々はそれに合わせて仕事をするが、そのようなコンテキストの判断は困難である。

以上を総合すると、ゴール設定から課題の見きわめや、さらにその分析、問題の投げ込みなど、いわゆる課題解決プロセスはボトルネックだらけであるということはやはり強く認識しておく必要があるのではないかと私は思っている。AI が人間かのように議論されていることは非常に危ない。

基礎知覚も、先ほど申し上げたように欠落しているし、そもそも感情がないので理性もない。予測は非常に強いが、問題の定義や、気付きなどは全くないわけで、意思決定もないという意味では人間を代替するというよりは補完するものであって、やはり我々が普通に理解しているような **problem solving** マシンではないというのは強く認識しておく必要があるのではないか。

日本人の場合、ドラえもんや鉄腕アトムのようなものばかり見過ぎている可能性がある(笑)。私は無理だとは言わないし、科学者であるのでこの未来を信じたいが、今のところ未来があるようには思えないというぐらい遠く感じられる。

AI が人間を代替するという議論はやはりおかしくて、AI は幅広く人間をアシストしていくものである。ここから先の人間は、すべて自分や周りの人の経験と知恵だけで頑張る人と、圧倒的なコンピューティングパワーやデータの力を使う人との戦いである。これは人と人の戦いである。ブルドーザーで穴を掘る人と、手で掘っている人の違いのようなもので、当然ブルドーザーを使いたいわけである。そのブルドーザーがデータであり、コンピューティングパワーであり、アルゴリズムであるというのが本当の世界観だと思う。後者の世界観でどうやって何をやるのだという形だと思う。情報の掘削力が上がっているということで、人間とは何かというのがかなり赤裸々になっているのではないかと思う。「識別」、

「予測」、「実行」がデータとAIの力で上がっていくことで、人間が総合的に見立てる、組織を率いるなど、そういったことをやっていくのだろうと思うわけである。

ビジネス・経営への意味合い（資料 p.77～90）

ビジネスや経営はどうなっていくかであるが、そもそもマネジメントはどう変わるのか。ここにおられる方々や私もマネジメント的な立場にあるが、相当大きな変化があると思う。データとキカイ（AI）に何を任せて、一人人間はどこで価値を生み出すべきかを見極める。これを間違えると、機械で一瞬で終わるものを人間にやらせたために企業がつぶれる可能性があるという、非常に重要なクエスチョンになってしまうのではないかと思っている。これまで経営の資源はヒト・モノ・カネと言っていた。たしかにいずれも大事であるが、モノやカネすら、こういうデータやキカイ（キカイというのはAIやロボットをイメージしている）に半ば依存することになることを考えると、ヒト・データ・キカイの時代に向かっているのではないかと思っている。

そのアセットでモノ・カネ系とデータ・キカイ系ということで、いろいろな産業分野をプロットしてみると、資料 p.80 のように、いわゆる昔からある incumbent というかオールドエコノミーが左上にあり、私が働いているデータ系の産業は右下、ニューエコノミーと言われているビジネスは右上にある。モノ・カネの重要性が消えるわけではないことを考えれば、この向かっているところは明らかにここである。すべてここに向かっている。谷口委員の自動運転車のようなものはまさにここにあるし、Elon Musk がテスラはメーカーであって Google とは違うのだと強く言っているのも、こちらに来ているのだという意識が強いのと、Uber などは半ばこの遺伝子を持っている。ただ、私の周りにいるような人は、正直言ってアトムに対する嫌悪感がある。スピードが 5 桁ぐらい遅いのでやっていられないと思っている。モノ・カネ側の人たちはこの人たちでデータ・キカイ側を胡散臭いと思っている。だから、第 3 種の人類が来るのを待っているわけである。第 3 種の人類として谷口委員はおられるわけである。ここにウワッと人が入ってくるのが未来を作っていくという感じだと思う。こちら側に移住できない会社は、特にこの青いオールドエコノミー系は衰退していくと思う。こちら側の移住ゲームがマクロ的に起きているのではないだろうかというのが私の感覚である。

では、ほかの業務はどうかというと、今のところ、マネジメントは半ばエクセキューション管理にかなり時間を使っているが、そこをかなり自動化できるようになれば、何に注目すべきかというようなことを投げ込むのがかなり中心の業務になるのではないと思う。

通常のデシジョンメイキングよりも、異常値対応といったものをどんどん突っ込んでいく。異常値であるなどというのは、普通のスピードの 100 万倍ぐらいの速さで上がってくる。今まで気付かなかったものが全部上がってくる中で、どれが本物か、どうするのかなどというのはパターン化できず、我々が考えていくしかないので、これは非常に重要な仕事になっていくのではないかと。

リスク管理というと、会社の偉い方々が本社の記者会見で頭を下げているイメージであ

るが、それは消えていく可能性が高い。人間はこういうことに関わらない。むしろ完全にブラックボックス化していくことの危険性のほうがかなり高いと考える。リーマン・ブラザーズが破綻したときも、AI とは言わないが、アルゴリズムの化け物のようなもので動いていて、それが1日5%以上動くとき既定のプログラムが動かなくなるようなことを誰も知らなかった。過去を振り返ってみると、チーフトレーダークラスが自分たちの使っているモデルが何に依存していて、どういう限界があるかを分かっていなかったことが理由だったということがいろいろ赤裸々になってきている。同じように、Deep Learning ベースだと変数の固まりで誰にも分らないという問題がある。それが相互に依存し合っていると、完全な systemic failure (システムの全停止) が起きる可能性がないとは言えず、非常にリスクを伴うだろうと思う。その間にあるような変数を半ばダイヤルのようにコントロールするようなことが我々にとって必要になってくるのではないかと思っている。

またソフトスキルは今非常に重要であるし今後も重要なわけで、AI やデータの指示で人間が動く時代は若干遠いと思う。コンテキストを与えるのは我々にとって結構大事なことで、それがないと人は動かないので、コンテキストライズしてつないでいくというのが非常に重要な仕事になってくるのではないかと感じている。

ビジネスはというと、ここまでの話から見て明らかであるが、一定規模以上の組織ではAI やデータの取り組みはもう不可欠になることはほぼ明らかで、これをやらなかったら単純につぶれていくというか、一瞬でやる人と10時間掛けてやっている人の戦いのようなになるので、そうなるだろう。加えて、意思決定の質が劇的に良くなることはほぼ確実である。それは普通の情報抽出やパターンから外れたものの検出が劇的に良くなるからである。ヤブオクの悪い人のパターンをたたき出すのは、要は異常値をたたき出しているわけである。丸山委員がおられる PFN (Preferred Networks) のデータなどの話を聞くと、4週間ぐらい前の誰も気付かない機械の振動などのデータを見て、この機械は3週間後にだめになるというのが分かってしまうわけである。

また、PDCA のサイクルはつながりが切れてしまい、人間が介入するのは非常に難しく、全体をどう回すかを人間がチューニングする時代になる。そこで PDCA の P だけをやるのだと言っても、速過ぎて分からないし、壊れていくわけで、かなり異質なタイプのサービスの回し方になっていくだろうと思う。今のデジタルマーケティングの世界はまさにそうである。

先ほどの AI の性質を見ると明らかなように、データホルダーとデータホルダー、データホルダーと技術プラットフォームの連携のようなものは間違いなくここから勝負になっていって、集合知的な AI をどう作れるかのゲームになっていくと思う。日本でも海外でも続々とこのようなタイプのゲームが出てきている。

自動化できるものは自動化されてしまうので、結局そこから残った引き算の部分のサービス提供や、デザインなどのヒューマンタッチ的なところが今まで以上に重要になっていく可能性が非常に高い。しかも、差別化ができにくくなっていく可能性が高い。

今後、データ・キカイ時代のヒエラルキーができてくると思うが、どちらも持たない会

社より両方持っている会社が当然上なわけである。問題はデータだけ持っている会社と技術だけを持っている会社のどちらが力を持つかであるが、私は確信を持っていて、技術プラットフォームだけを持っている会社よりも、データを持つ会社のほうが力を持つと思う。なぜなら、データを入れないと AI は動かないことと、結局データを持つような場、データが生まれる場から一番データの価値が生まれてくるということで、それはかなり明らかな流れかなと思う。技術そのものは発展のためには概ねオープン化されざるを得ない。Google のテストフローなど、いろいろなものがあるがオープン化されているのはかなりこれに近い話だと思っていて、これらのものは一般の人に使ってもらって改良していく必要がある。昔の UNIX のような世界になっていくのではないかなと思う。

日本では丸山委員がおられる PFN とトヨタ、あるいは PFN とファナックのように、アメリカでは IBM の WATSON が全米最高レベルの医療機関である Mayo Clinic や医大などで、年間 8,000 万例の医療データのほぼすべてを吸い上げている Epic のデータを丸ごと吸い上げて、診断の完全自動化を図ろうとしている。全く同じような動きが Google の DeepMind であり、DeepMind は NIH と NHS それぞれと非常に特殊な契約を結んで情報をほとんど吸い上げている。医療の付加価値はよく cure / care / prevention と言うが、本当の価値は診断(diagnosis)から来ているわけである。頭が痛いと言う人が脳出血なのか、あるいは脳のがんなのか、判断を間違えたらその人は死んでしまうわけで、正しい判断がやはり勝負であるが、こういったことをこれほど巨大な産業セクターで AI がいかにも効きそうなところから取り込もうという動きが起きているのは、まさにこういう話だと思う。

経産省の産構審で議論している運転の自動化の話であるが、レベル 4 の完全自動運転になった瞬間に産業構造が根底から変わるだろうと言われている(資料 p.86)。ここまでは車を保有することが大事であるし、メーカーの世界であるが、ここから先は多分サービス業の世界になってしまって、多分乗っているだけの人が出てくると、周りの空間をどうするのかなど、あらゆる全く違ったサービスが爆発的に生まれる可能性が高い。このレベル 4 的なものがどんどん出てくることで産業の質がかなり変わる可能性が高いと思っている。

産業革命のときをしてみると、最初に新しい内燃機関や電気機械などが出てきた時代と、それらをベースに高度な応用が進んでいった時代と、さらにそれをベースにもっと深い系ができてきた時代の 3 段階ぐらいにざっくりとメタ化して考えることができるのではないかなと思っている。これと同じ視点がデータ化、AI の世界でも結構大事だと思っていて、フェーズ 1、2、3 とすると、今騒がれているデータ・キカイ化が進展というのはフェーズ 1 の話で、これは先ほど見た通り行くしかない。no regret。そうしないと、消えていくか、もしくは版画で人の顔を描いているような世界になってしまう。別に浮世絵で顔を描くことは否定しないが、誰もやらない。版画の 1 億倍ぐらい写真が使われている。そのようにこれは行く。ただ、ここから 10 年ぐらいのフェーズ 1 のような話に目をとらわれたらまずい。やはりモーターなどを与件に VTR ができてきたように、このデータなり AI というのが与件になって何かが出てくるというフェーズ 2、それらが組み合わさって全体の系として何をやっていくというフェーズ 3 のところをどうやって考えていくかがやはり勝負なので

はないかと思っている。というのも、日本はそもそも産業革命のフェーズ 1 のときにはちよんまげを結っていて何もやっていなかったわけである。この国は寝ていたわけで、このときはチャンバラや田植えをしていた。この時代に我々は参画していないままフェーズ 2、3 に変わり、深く考えて勝つ。フェーズ 1 のときのワットやファラデーではない。このとき金持ちになったのは、少なくともこのフェーズ 1 のところにいたエジソンのような人で、その少し後にやってきた松下幸之助翁のような人たちなわけである。したがって、やはりこちら側をどう作っていくかというところを集中的に考えるのが結構勝負なのではないかと思っている。

また、産業群が先ほど見た通り全く違う 2 つの群に分かれているので、全然違う立ち位置で議論すべきであり、このデータ・キカイ側の産業とモノ・カネ側は全然違うから、この視点も要るだろうと思っている。人材要件が変わってくるわけである。

成功要件と日本の現状 (資料 p.92~105)

最後に、成功要件と日本の現状の話を少ししてみたい。

データの利活用は、先ほどの収集、処理、出力の 3 ステップで、この枠組みで見たときに 3 つ大きな要件があるのではないかと思っている。デバイスや領域を超えたマルチビッグデータの利活用ができるのがまず基本的な要件であることはほぼ確実である。次に圧倒的なデータ処理力がある。これは技術もあるが、そもそもデータセンターのようなものである。さらに、質量ともに世界レベルの情報系サイエンティストとエンジニアリング、この三位一体で基本的にはやっていくのが、ここからの変革におけるある種のエンジンだと思う。この三位一体が必要だと、この産業にいる私としては日々感じているが、例えばデータという視点で見ると、検索、ポータルはヤフーが日本では多いが、Google は 1 桁多い。e コマースでは楽天が日本では強いが、Amazon は 1 桁多い。SNS は半ば崩壊してしまっている。ミクシィが一応日本では強いが、Facebook は 2 桁多い。チャットも LINE が強そうに見えるが、日本は 3,000 万しかいないから、やはり桁が違う。非常に巨大なデータを持っている連中と我々は戦っている。これは、やはりどうしても意識せざるを得ないところがあるかと思っている。

また、カオスマップと言われているバリューチェーンの中にプレーヤーを突っ込んだものを見ると、メッセージはクリアで、この中に全然日本の匂いがしない。ヨーロッパもない。ほとんどアメリカとイスラエルである。日本が強い産業の中身をここまで食い荒らされたことがあったのだろうかというぐらい結構危険な状態で、由々しき事態である。

ちなみに Google は AI 系の買収をものすごい勢いでやっていて、その幾つかを最近また売却しようとしているが、2 年半前の 12 月頃は 1 週間ぐらいで AI・ロボット系の 8 社を買ったりしていた。そこに日本のシャープなどがいたりするわけで、人間の脳に当たる部分、手に当たる部分、足も全部買ってしまえば、人も全部買えば、こういうゲームが行われている。これはどうなのだというのが一つ。

次に、情報処理のところであるが、Amazon 等の海外向けのクラウドに国内外ともに依

存しているのは非常に良くないと思っている。これはある種の安全保障上のリスクもあるし、競争力の源泉でもあると思うわけで、こういったものが外に漏れ出ているのはどうなのだろうか。非常に危惧を抱いている。

また、投資規模も、よく誤解してこれが AI などと言っているが、そのようなものはやはり情報投資規模で、この人は AI だったらこのぐらいしか無いはずであるが、ただ、とんでもない額でやっている。ざっくり年間 8,000 億円、1 兆円という桁でやるので、しゃれになっていないとは言える。

また、人材が足りない。日本というのは無意識にアメリカの次に人材がいると思いがちであるが、そのようなことはなく、中国とインドに負けてしまっている。4 番目である。しかも、この中身というのは、昔からの SI 屋のような人たちが多く、今必要なタイプのビッグデータ系の時代人材ではない人が中心になるのも頭の痛いところである。

この一つの背景は、昨年丸山先生や文科省の連中と散々議論したが、そもそもこの国は理工系の大卒の数が足りない。その実数は、人口 5,000 万の韓国より日本は何と毎年 10 万人も少ない。このような状態でどうやって勝つのか。大卒に占める理工系の率は 23% しかない。この国で理系だと言うと半ば変態になってしまう (笑)。大学院で博士号など、頭が大丈夫かと言われてしまう。ドイツや韓国は大卒に占める理工系の割合が 6 割を保持している。こういうところを見ても、人材のベースがそもそも大事である。ドイツは一回沈んでいたが、この 10 年ぐらいで V 字回復している。明らかにこちら側が成長産業分野だということで人が寄ってきて、きちんとそれが育っているわけである。なぜ日本はこうなっているのか、どういう問題があるのか。これは結構まずい問題なのではないかと思っている。

データは若干古いが、日本でトレーニングを受けた人の数が足りないのは、見れば明らかである (資料 p.99)。日本人はアメリカ人は理数系に弱いと思いがちだが、そのようなことはない。100 人あたり人数を見てもアメリカは日本の 3 倍ぐらいいるわけである。やはり日本は何か教育が間違っている。人を育てる方針を変えないと、本当にまずいことになると思う。

実際私はアメリカの少しまともな街のコンビニに行くと驚くのは、普通に『サイエンティフィックアメリカン』などが置いてあることである。日本で『日経サイエンス』が置かれているコンビニなど、私は見たことがない。大学生協ならある。

長い間採用をやって、今も幹部トレーニングをやっているが、優秀な大学から 50 倍、100 倍でとつても、そもそも問題を定義できない、指数と実数の使い分けができない、数字を並べることと分析の違いが分かっていない、平均を鵜呑みにする、このような学生が大半である。だから、会社でゼロベースでたたき直さなければいけない。高等教育を受けた学生が高等教育を受けていないというようなおかしい状態で大学を出てこられては困ると思いつつながら、良い人を採り放題な状態でもこうなってしまう。

サイエンスの専門家層はそもそもいないし、どこにいるのか分からない。いても、そもそも社会に関心がないなど、結構厄介な状態である。強化だけでは駄目で、マインドシフトを変えないといけない。オタクではない、世界を変えるようなギークやハッカーが必要

である。シリコンバレーを見れば、Jerry Yangにしても、Larry PageやSergey Brinにしても、Elon Muskにしても、Andy Rubinにしる、皆テックギークである。こういう人たちが世界を変えてきているわけで、ビリオネアになっていく。これが世界の流れである。

先々週、NII 所長の喜連川先生から東大生に喝を入れてよと言われて、東大の情報系の院で授業をしてきた。院生からの質問の一つ「このまま情報系で Ph.D.を取るのがいいか、MBAに行くのか、どちらが良いと思うか？」には「そのようなものは Ph.D.に決まっている、何を言っているんだ、君たちは」と。また、「今、日本社会は我々のことを本当に必要としているのか」という質問には、「君ら以上に必要とされている人はいないが、君たちの価値を理解してくれる人がいない会社はつぶれるから行かなくていい。君らの価値が分かる人のところに行けばいい。そこが駄目だったら自分で仕事を始め、バイアウトすれば生きていける」と答えた。それぐらい不安に満ちている。東大の情報系の院生はスタンフォードと同じぐらいセレクションされているはずである。何でそうなるのだと非常に不安を感じて帰ってきた。彼らが世界を変えて、ビリオネアになったらどうやって生きていこうかと妄想しながら世界を変えていってこれないといけないうところで、何かおかしくなっている。

エンジニアリング層も、先ほども言ったが、プログラマーや Sler 的な人たちが中心で、研究と開発のギャップを越えられる人は足りていない。ビッグデータ処理などを回せる人が全く足りていなくて、データサイエンスとつなげられる人も足りていないという意味で、やはり由々しい問題である。

もっと問題なのは、ここにおられる委員の方は例外だが、私が見ている範囲では、そもそも今ビッグウェーブが来ているという感覚が全くない。今はエジソンのような人物が生まれてくる時代なわけである。松下幸之助や盛田昭夫になれるという時代に生きている感じが全くない。我々が持っているデバイスなどは、海外の力によって信じられないほど進化しているが、そのような状態でいいのか。アーキテクト的な人もいないし、その人たちが生き延びるためにはそもそもスキルを renew しなければいけないが、どうやったらよいか分からないということで、非常に厄介な問題である。

必要になる取り組み（案）（資料 p.107～123）

では、何をやっていったらいいのか。まずデータ系については、データの利活用がしやすいような状況に持ち込むのが重大だと思っている。後追いでは駄目である。実験的な試みがどんどんできるような必要があり、若い人たちが思い付いたことをどんどんできないといけない。これは駄目だ、これは申請しろなどと言っていたら、彼らの人生の 1 年は非常に長いので諦めて次のことをやり始める。データ利活用規制は本当に危ない。また、このようなことは無視しているアメリカや中国企業の問題がある。今は個別対応で、申請があったら国家戦略的にそこだけ認めるというようなホワイトリスト方式でやっていっている。私は、これでは駄目だと思っている。基本的には思い付いた人が人道に反さない限りやれるという条件を持ち込む必要があり、ホワイト方式ではなくブラックリスト方式でな

いといけない。刑法に触れるようなこと以外は全部 OK であるとしないうり変わらない。非常に patient な人だけが頑張れる。普通の若者は無理である。突破力と patient は別の才能で、これを併せ持っている人はいないので、ほとんどあり得ないことを期待していても仕方がない。このような確変モードのときには、思い付いた人がやれるようにしてあげないといけない。これは非常に重要な問題だと思っていて、私は特区を作ってくれと日々言っている。例えば劇的な復興が必要な東北や、あるいは九州、そして中国から身を守らなければいけない沖縄などは全部特区にさせていただいて、人を傷つけない限り OK というぐらいにしてほしいとやらないと、いつまでも問題が解決しない。それがやはり組織なり社会の自然治癒力を上げるのではないかと思っていて、これが結構重要なポイントである。アイデアを思い付いてから申請しても間に合わない。

それから、データ処理力の問題も結構大きいと思っている。データ自体が海外に流れていっている。一番大きい理由は、実はコストである。日本国内の DC は、世界的に見ても非常にコストが高い。実際データ処理の価値を生んでしまうわけで、データ処理のコストが高いということは、競争力を失うことである。当社はアメリカのワシントン州に DC を持っているが業務用の電気代が全然違う。日本の 6 分の 1、10 分の 1 である。10 倍も処理コストが違ったら、勝てるわけがない。しかも、これはアメリカであるから、中国やロシアと対抗したら 50 倍くらいの差がある可能性が高い。レナ川の水力発電ならどう考えてもただである。つなげば幾らでも電力がやってくる。そういう連中と我々は戦っている。黒竜江か何かよく分からないパワーで動いている。競争力が足りない状態をどうにかしないといけない。DC に回す補助金などは全部無税にするなど何かやってくれないと勝てない。

人材の話は、先ほど見たように非常に情けない状態であるが、今必要なのは基礎研究にしか関心がない人ではなく、この非常にエキサイティングな時代にエキサイトする人であり、統計だけの専門家ではなく情報科学的な知恵と技を課題解決に使う人であり、仕様書通りに回している SE などではなく、課題を俯瞰してきちんと実験環境からデータ処理ができるような人たちが必要なわけである。

小中高教育でネットを利用するなど、ネット業界の人たちが頑張ってきたので、触れることは本来問題ではない。そうではなく、ネットやアプリを作る授業を行って欲しい。大学教育で、高専等も含めてであるが、理工系の学生をとにかく増やして、現代の基礎教養としてのデータ分析教育はやって欲しいと思っている。また、国家的プロジェクトをやることも重要で、これは軍事費等も使ってリーダー層を作り込まないといけない。加えて IT エンジニアリングの再教育と、1,000 万単位でいるミドル・マネジメント層をどうやってリエデュケーションするかが重要な問題で、そうしないとこの人たちは食いつぶれてしまう。

リベラルアーツは、何となく日本ではゆがんだ意味になってしまっているが、これは 2000 年以上前のギリシア時代にどうやって奴隷ではなく自由人になって生きていくかという基礎的な技能だったと思う。それは中世においては弁論術なり、修辞学、数学、音楽だったわけである。では現代ではどうかと言うと母国語できちんと論理的に議論して明確に意思を伝えられ、英語、今は中国語も入りつつある世界語で情報収集などがある程度できる。

さらに問題解決能力。ここまでが、過去 10 年前までのリベラルアーツだったと思う。加えて今は、明らかにこのデータ時代に即したデータリテラシーと言うべき分析的、データドリブンな思考力と基本的な知見が追加で入ってきている。ここまでセットでやって世の中に出してあげないと、食いつぶされる可能性がある、この大きな波に乗れない可能性があるということで、この母国語、世界語、問題解決能力、データリテラシーの 4 つのセットでやってほしいと、かねがね文科省などには言い続けている。

データを解き放つために必要なスキルは、データサイエンス（情報処理）、エンジニアリング（実装する力）、ビジネス力（実装する力、それに実問題をきちんとつなげていく力）、この 3 つが必要である。これを三位一体でやるのが、RU11（リサーチユニバーシティの 11 校）、総研大、北陸先端大学、奈良先端大学、OIST や、その他国立大である。また、高専や進学校のすべての理系の学生にこれをやるぐらいのことをしないと人材は足りない。一気にやって人を育てる流れを起こしていったらどうだろうかと思っている。

リーダー層の話では、アメリカと日本で人口が 2.4 倍しか違わないが、日本とアメリカの科学技術予算を見ると、日本は 4 分の 1 しかない。これでは勝てない。すべて人口比にしてほしいというのが私の主張で、年間 600 億ドル程度出してほしいと訴えている。

日本は 90 年代まで行け行けだったので、辻井先生をはじめとする第五世代コンピュータの人たちが日本のリーダー層を作ってきたわけであるが、ここからは何もやっていない。アメリカは、その後ストラテジーコンピューティングイニシアチブをすぐ真似してやったり、ヒトゲノムに何千億も突っ込んだり、Decade of Brain など 1,000 億円級のプロジェクトをたくさんやっているわけである。日本には一応先進国としての国の格があるのだから、アメリカだけに依存したり、中国に走らせておくのではなく、きちんと何かやってほしい。そうでないと人が育たないと騒ぎ立てて、ようやく AI やビッグデータだとかで少し予算が付いたが、辻井先生のお話を聞くと真水が年間 15 億しかないとか、桁が 2 つ足りないのではないかと。真水を千億欲しいというのが私たちの主張である。だが、何もやらないよりよいので良かったが、それでは駄目である。

このままでは利活用が阻害されて、DC を回すにもカネが掛かる。データセンターが枯渇していれば、日本で何かやっている人たちの競争力がなくなっていき、さらに米国企業や中国企業にデータが集中し、日系企業だけが IT で飯を食えなくなっていく。負のサイクルが回り始めて、欧米に加えて日本企業も国外脱出を始めて、人材流出も加速し、税収も減って、都市や地方は荒廃し、競争力を失うという悪夢のサイクルに入っていくというのが私たちの主張である。そうではなく利活用を促進して、世界中からも日本に人がたくさん来たいというぐらいにしておかなければいけない。

日本には山ほどデータがあるとなれば、皆やって来る。そうすると、競争力が上がっていったら、アメリカへの集中も止まり、好循環に入って、日本はもうハブになって、80 年代で終わったはずの流れがもう一回戻っていく。負のサイクルではなく、正のサイクル、not this but this でやっていただかないと、私たちの未来はどうなってしまうのかと、老後が不安で生きていけない。やれる限り頑張るが、このままではまずい、非常に良くない。

才能と情熱を今解き放たなければいけない。解き放って圧倒的に活用しやすい国を作っていけたらよい。案外、難しそうであり難しくないなので、がんがんやっていく解き放ち方を研究するのが私たちに課せられていることではないかと思っている。