

人と AI の共生がもたらす社会変容

栗原 聡¹ (慶應義塾大学)

要 旨

2020 年以降、大規模言語モデル (LLM) を中心とする生成 AI の実用化により、AI は再び大きな注目を集めている。AI は当初、道具としての効率化を目的としていたが、現在は創造性を高めるツールとしても期待される。AI は新しいアイデアをゼロから生み出すことはできないが、人が既存の知識を組み合わせることで新たな価値を創造する点で有用である。今後、AI はさらに進化し、いずれは AI 自体がノーベル賞級の発見をする可能性も期待されるが、一方、制御不能にならないよう慎重な対応が必要である。AI 研究は 70 年を迎え、道具型 AI の時代が終わりを迎えつつあり、自律汎用型 AI の時代にシフトしつつある。自律汎用型 AI は状況を理解し、適切な行動を自ら選択できる。そのために、動作するための目的を持ち、人が細かく指示を出さずとも、AI が自ら状況に応じた判断をすることが求められる。

しかし、それ以前に現在においてインターネットは膨大な情報であふれ、生成 AI による偽情報も増えている中、人が直接アクセスすべきではない媒体となりつつある。この問題に対処するため、情報の出所を明確にする Originator Profile などの技術が提案されているが、それでも限界がある。根本的な問題は、人間の情報処理能力が物理的なモノと異なる情報には適していないことにある。物理的なモノは形や質感などの「身体性」を持ち、それが行動を促すが、情報にはそのような身体性がないため、文脈を理解することが難しい。インターネット普及以前、情報拡散は局所的であり、人々は情報の背景や文脈を把握しての対応ができたが、現在では膨大な情報に圧倒され、情報に含まれるデマやフェイクを識別することがどんどん難しくなりつつある。この状況で、自律型 AI が適切な情報を収集し、多様性を保ちつつイノベーションを促す情報を選択する未来が期待される。

自律型 AI は、個人の利益に配慮しつつも社会全体の利益を考慮した判断ができる可能性があるが、AI の判断を人が信頼するためには、説明可能性が重要である。また、今後において因果推論技術などの発展により、AI の判断に対する信頼も高まることが期待される。人が AI の判断を受け入れるには時間がかかるが、成功体験を積み重ねることで信頼が構築され、最終的には複雑な問題を AI に委ねることができるようになると考えられる。AI に判断を委ねることは、AI による支配ではなく信頼の上での人の選択であり、人と AI が共生する社会において、AI は人間の創造力を高めるパートナーとなり得る。

キーワード：生成 AI、自律型汎用 AI、創造性の向上、説明可能性と信頼性、情報過多

1. 創造することとは？

ダートマス会議を AI 研究の開始とすれば、そろそろ 70 年が経過しようとしている。2

¹ 慶應義塾大学理工学部教授／人工知能学会会長

回のブームを経て現在は3回目の真っ只中にあり、それでも2000年くらいからの Deep Learning への加熱が落ち着いてきた矢先、2020年に入り ChatGPT を筆頭に生成 AI が開発され、加熱ぶりが再燃するどころかひたすら盛り上がっている。過去2回のブームはそれぞれ失敗に終わったと教科書には書かれているが、現在から振り返ると、初回のブームはコンピュータの高性能化やインターネットの拡充、そして、それらをインフラとしての Deep Learning の実用化をもって、当時提案された様々な手法が実用レベルに到達しており、紆余曲折あったものの成功に終わったと言えるのではないかと。また、2回目のブーム以降、AI の最大の壁は我々の常識や、五感も駆使しての状況把握能力などであったが、ChatGPT レベルの大規模な LLM の開発の成功をもって、いよいよ常識を取り込み、最大の壁も突破しつつある段階に至っていると言えよう。そして、LLM をエージェント化して利用する流れや、ロボットに LLM を搭載する研究が活性化しつつある現在、70年に及ぶ道具型 AI の時代が終わろうとしているとも言えるのだ。道具への信頼は動作の正確性にあり、これまでのテクノロジーの目的は主として効率化にあった。しかし、効率化自体は創造的な行為ではない。創造的な作業は今も人の役割である。リソースに乏しく、国際的な生産力も低下しつつある日本においてはイノベーションの多産による国力増強のみが唯一の打開策であるとすると、AI の創造的な作業への活用こそ積極的に推進すべきである。

しかし、日本における生成 AI 活用の度合いは海外に比べて低く、DX化においても総じて進行が遅れている。DX化の遅れについては、DX化せずともしっかり仕事のできた日本人としての気質が裏目に出ているのかもしれないが、その結果として IT リテラシーが総じて低く、最先端テクノロジーを活用できなければ、効率化もままならず、AI のイノベーションへの活用以前の段階である。

AI を実際に活用することもなく周りからの声に影響されるだけにおいて、「AI は危険である」とか、「生成 AI は万能ツール」のような漠然とした感覚を多くの人々が持っているように思える。「AI が人のようにゼロから創造する能力を持ったら危険である」といった声もよく耳にする。

そもそも、現在の AI は新たなアイデアや価値をゼロから創造することはできない。実は冷静に考えれば、我々人も AI と同じである。人間においても、これまで自ら経験して学んだことや、学校等で教えられた知識以外の新たなアイデアがいきなりゼロから生み出されることはない。しかし、我々は自分の頭にあるアイデア同士を組み合わせることで新たなアイデアを創造することができるし、ブレストにて自分にはないアイデアが相手からもたらされたことで新たなアイデアを生み出すといった、知識を組み合わせることで新たな価値を創造することができる。研究において、その研究分野を推し進める際の表現に「巨人の肩に乗る」という言い回しがある。1人の研究者がなし得る業績は決して大きいものではない。しかし、1つの研究が新たな研究の基礎となり、その積み重ねとして巨人の高さほどに研究分野は成長する。よって、新たな研究はその巨人の肩に乗ることでさらに先に進むことができるという意味である。

では、ノーベル賞級に偉大な発見といった、ゼロからと言える類いの大きなイノベーションはどのようにして起こせるのであろうか？ それは、失敗や偶発的イベントなどにより、「意図せず」想定外の知識やアイデア飛び込み、自らが持つ知識を組み合わせた時に創発されるのだと言えよう。まさに失敗は成功のもとなのだ。

その意味では、プレストもだいたいにおいて既存の相手とのやりとりであることが多く、お互いの考えそうなことは大方認識されていることから、想定外の知識は得られにくいかもしれない。敢えて新たなメンバーを入れることが重要というわけだ。かといって、極端に無作為に適当な人を集めればよい、というわけにもいかないのは当たり前である。すでに過去の遺物となってしまうが、企業でも大学でも「たばこ部屋」なるものがあった。喫煙のためではあるが、所属や部局に関わらず人が集まる場所としては、実はイノベーションが生まれる可能性が高い、皮肉なことに魅力ある場でもあったのだ。現在においてはこのような場もなくなり、後述するが、ネット空間での多様性も低くなりつつあるなど、イノベティブな社会に突き進むべきこととは逆の方向に突き進んでしまっている。

これに対して、ChatGPT 級の LLM の知識は膨大であり、一人一人が持つ知識を大きく凌駕する。つまり、各人にとって十分に想定外の知識が過分に詰め込まれていることを意味する。LLM を壁打ちにしてアイデアを練る、ということよく耳にするがまさに正しい。AI をプレスト相手とすることで想定外の知識が AI から提供されることによる多産なイノベーションの可能性が期待できる。ただし、効果的な壁打ちのためにはプロンプトを工夫する必要があるものの、重要なことはプロンプトで AI に多様な役割や状況設定を与えることで、いかにしてランダムとまではいかないものの意外性のある回答を引き出せるかであろう。

しかし、どれだけのイノベーションを起こせるかどうかは、結局は我々の能力次第であることには変わりはない。我々の繋ぐ能力次第である。多様な物の見方ができる能力や、状況を理解する能力、そして文脈を正確に把握できる理解といったごく当たり前の能力である。しかしその当たり前の能力が、ネット依存度が急激に高まっている現代人、特に若年層において低下しつつあるように見える。そして、この能力は決して学校で教えてもらう類いのものだけではなく、日常生活における友人や家族、他人とのやりとりにおいて自然と育まれるものであろう。現実においては多様な人々と対峙する必要があることでの社会性を身に付けられるはずが、ネット空間では嫌いな人とは陽に関係を立ち切り、好きな人同士で群れを作るエコーチャンバー（フィルターバブル）という現象が顕著となり、これも繋ぐ能力を育むにはマイナス効果である。

繋ぐ能力は、訓練することで大人になってからでもそれなりに高めることはできるのかもしれないが、やはり初等教育の時期に繋ぐ能力を身に付けることが大切であろう。Youtube で多様なコンテンツを見たところで視覚情報のみである。実世界は複雑であり、この複雑な世界との能動的なインタラクションこそが繋ぐ能力を育む土台であるはずが、ネット空間には物理空間のような複雑性は高い。これに対して、友達と公園で遊ぶとか、砂場で遊ぶとか、視覚・聴覚・触覚・嗅覚・味覚のフル稼働である。これが脳内における思考のためのネットワークを成長させることに繋がる。本を読み、長文を理解する能力や行間や文脈を理解する能力も重要な繋ぐ能力であろう。そのような能力があつてこそ、自分の考えを論理的に主張することができる。

「答え」より「答えを導出するプロセス」が大事であるということを知ったことがあると思うが、これも繋ぐ能力を育むことになる。問題を解決する方法は一つではなく、複数の方法を考えると柔軟な発想ができることも重要だ。そして、物事を俯瞰的に見る能力も重要である。そして、これらの能力は基本的に身体動作を伴う作業や人とのコミュニケーションなど、社会の一員として成長することで獲得される。つまりは、初等教育におけるの端末を

与えての過度な情報教育は、繋ぐ能力の成長にとってはマイナスの効果が大いかもしれない。

肝心の繋ぐ能力が低下すれば、折角のAIとの壁打ちの効果が低下してしまう。いわゆるイノベーターと呼ばれる人々は繋ぐ能力が高いということだ。一見関係がないと思われるアイデア同士を繋ぐことでの価値の創造ができるという意味である。

筆者は、トップクリエイターがAIを活用して新しいマンガを制作するTEZUKA2020プロジェクトを2020年に実施し、「ぱいどん」という新作を制作したことがある。2020年はまだ生成AIが登場する以前であり、我々が開発したAIにより、複数のプロット（あらすじ）とキャラクターの顔の画像を事前に生成し、その中からクリエイターが使いそうなものを選択し、作品を作り上げるための素案として活用するというやり方をした。ChatGPTを使って生成するような、読み物としても成立するレベルのストーリーが生成できるわけではなかったものの、クリエイターの発想のためのたたき台となる素案を生成することはできた。

まず、素案を大量に生成し、制作会議に持ち込むこととなった。素案の中には我々から見てもストーリー展開に一貫性があり、面白いと思えるものから、話の流れに一貫性がなく使い物にならないレベルのものまで様々であった。ここで驚かされたのがクリエイターの反応だったのだ。我々が面白いと思った素案や一貫性があると思った素案は即落選となってしまったのだ。理由は「面白くない」とか「この程度なら誰でも思い付ける」ということだった。そして、我々がデタラメと思った素案に対して「これは面白い」と高く評価したのである。作品「ぱいどん」として完成したマンガの素案となったストーリーのそのような我々としては低い評価の素案だったのである。

気づかされたこと、それは前述したようにクリエイターが一般の素人よりもアイデアの種と種が離れていてもそれを繋ぐことで新たな発想ができるということだ。我々には離れすぎて繋ぐことができず、その場合はデタラメな素案としか思えないものが、クリエイターにとっては繋ぐことができ、さらに、より離れている種どうしを繋げた時こそ奇抜さや新鮮さを生み出せたということなのだ。

ということは、AIも繋ぐことでの新たなイノベーションを生み出す能力を持ち得ることは十分に考えられるとうことになる。しかも、膨大な知識空間においては繋ぎ方の数も膨大である。膨大な論文を読み込み知識を入れ込むことで、2050年にはAI自体がノーベル賞級の発見をすることを目指す取り組みがあるが、その可能性は高いのだと思う。

さらには、AIに対しての偶発的な知識の入力があることでのAIによるゼロからの発見的なイノベーションも可能となるのかもしれない。創造する作業においてはAIに追い越されるのは時間の問題なのだと思うが、加えて、AI自体を想定外に進化させることでもAIによる未知なるイノベーション創出の可能性すら考えられる。ただし、AIの進化については慎重であるべきである。いろいろ議論されているところであるが、AIが自らの目的を書き換えるなどとすることで、我々からの制御が完全に効かなくなる可能性があるからである。

AIの可能性について述べてきたが、我々人類はどうであろうか？ 地球沸騰化や、デマやフェイクが氾濫する日常、戦争はしているし、問題山積である。政治家の最後の決めセリフが「対話で解決」である。もちろん正しいしそうであるべきである。しかし、人類の自助での抜本的解決は極めて困難であることは自明のようにも感じる。あたふたしている間に

人類を乗せた舟自体が沈没しようとしている。このような状況に加速的に至らしめたのは明らかにテクノロジーであり、それは我々が求めたものである。であれば、救うためにもテクノロジーを活用しなければならない。可能性は AI にある。

2. AI は道具型から自律型へ

冒頭にて、そろそろ AI 研究が開始されてから 70 年となり、道具型 AI の研究開発の時代も終わると述べた。では、次は何なのかというと、それは自律型汎用 AI の時代である。

道具型 AI は人が明確な用途を決めて利用する。どのように使おうと人次第である。そして、人の創造力を超えた使い方はできない。ChatGPT は汎用 AI が実現されたとも評される。実に多様な種類の問いかけが可能である。ここでの汎用 AI は「汎用型の道具としての AI」という意味である。キャンプなどで使う十徳ナイフのようなものであることは先に述べた通りである。どの状況でどの道具を使うのかは人が決めなければならない。器用不器用の差がここで出る。汎用 AI の可能性は大きい、それは、汎用 AI を使い熟す人次第ということだ。

これに対して、自律型 AI は十徳ナイフ自体が状況を理解して、どの道具を使うかを十徳ナイフ自体が決めることができるようなものだ。ただし、状況を理解して使うべき道具を選択すること、すなわち、自律型 AI が状況を理解して、その場その場に適切な動作をするためには、現在の AI が持ち得ない重要な能力を付加する必要がある。能動的に動作することは、そのためのインセンティブ、つまりは動作する「目的」を AI が持っている必要があるのだ。自律型システムとしての我々生物の目的は当然であるが「生きること」である。

自律型 AI への目的の設定は設計者である人がするわけであるが、具体的に身近な設定であれば、家を綺麗にせよとか、仕事のサポートをせよ、といった抽象度の高いものとなる。AI は状況を理解し、人から与えられた抽象度の高い目的に対して適切な具体的な目的（サブ目的）を自ら設定してこれをベストなタイミングかつベストな方法で能動的に実行する能力が求められる。

人は楽をしたい生き物であり、いちいち細々と命令するのは面倒であり、AI が場の空気を読み適切な行動を自ら選択・実行できることは人にとっても理想的であろう。

自律型 AI の実現には状況を理解する能力を始め、周りの変化に即応的できるだけでなく（System1 と呼ばれる）、じっくり考える能力（System2 と呼ばれる）を適宜使い分ける能力も求められ、その実現にそう時間はかからないと思うが、すでに LLM を自律型のエージェントのように利用する方法がどんどん開発されつつある。LLM に汎用性があるからこそその応用である。エージェント同士を対話させることで新たなイノベーションに寄与するアイデアを人に提供する方法も極めて有用である。

ただし、汎用性があり、熟考すら可能とした ChatGPT であるが、ChatGPT の土台である Transformer は、単語の繋がり方の特徴を学ぶ仕掛けであり、それだけであるにも関わらずこれだけのことを可能としたことは驚きであるものの、万能とはさすがに言えない。高度な論理的思考や数学的な思考は Transformer 単体では難しい。また、System2 的な思考を System1 のような早い思考で可能としたことは間違いないものの、それは ChatGPT が持つ学習により獲得した知識の範囲においての熟考ではある。そうであっても ChatGPT が持つ知識量は絶大な大きさであるからとんでもない熟考ができることは間違いないが、

我々が何か熟考する時、頭にある知識のみを材料として熟考することはまずないはずだ。

我々は常にその場の状況を考慮しつつ早い思考や遅い思考を駆使して行動しているはずだ。常に目や耳といった五感から周りの画像や相手からの音声を取り込み、それらの情報と、自らの何らかの目的を達成すべくその時々における妥当な行動を能動的に選択しつつ活動しているのである。突発的なことが起こればそれまでの行動を瞬時に変更することすらそれなりにできてしまう。ChatGPT 単体ではこの適応的な思考ができないのである。あらかじめ与えられた知識のみを対象とした思考しかできないのであるから当然であろう。

AI 技術は加速的に進化している真っ最中であり、人は相手の気持ちを推し量ることができる、それを他者理解と呼ぶが、相手がどのようなことを考えているかをシミュレーションできる能力も必要であろう。さらに人は、他者理解において相手の思考をシミュレーションし、そのシミュレーションにおいて相手が自分をどう見ているかすらシミュレーションできる。ChatGPT は人のやりとりと同じレベルの流暢なやりとりができるので、つい人のような思考ができると勘違いしてしまうが、まだまだ限定された能力が実現されたに過ぎないのである。

そして、高い自律性と汎用性を兼ね備える AI の登場により「OMOTENASHI (おもてなし)」ができる AI の登場が現実味を帯びてくる。「おもてなし」とは、相手の状況を推察し、ベストなインタラクションをベストなタイミングで相手が行動を開始するよりも先にプロアクティブに実行できる能力と言える。

すると、道具型 AI としての生成 AI との壁打ちでは、プロンプトを工夫するなどしないと偶発性やセレンティビティを起こすことは容易ではなかった限界が、OMOTENASHI-AI では、AI が能動的に相手の状況を理解して、適切なアイデアを人に提供できるようになり、より効果的にイノベーションを起こすことができるようになるのではないかと。

実は、重要なことは人と自律型 AI とのやりとりを重ねる過程で「人の AI への信頼感が生まれること」にある。ここでの信頼感とは、道具型 AI に対する正確性とは異なり、自分から搾取しないという意味での信頼感である。必ずしも自分の思うとおりに行動してくれないかもしれないし、こちらからのお願いを聞いてくれないこともあるかもしれない。夏の公園、子供が暑い時にジュースを欲しがると、熱中症の懸念はない場合であるが、親がここでは取って買わないといったことは幼少期の記憶に誰しもあるかと思う。これは親としての子供の健康維持という、より長期的な目的に対して甘い飲み物の過剰摂取を避ける判断をしたからであり、子供のことを想っての行為である。子供が道具型 AI にジュースをお願いすれば、AI はジュースを調達するであろうし、拒否するような道具は使われなくなってしまうだけのことである。ここでの信頼感とは、自分のために動作しているのだという意味での安心感とも言える。

ただし、人の自律型 AI への高い信頼度がいきなり獲得されることはないであろうが、人の適応力の高さ故、ゆくゆくは AI の判断に対して人が主体的に同意しての AI の判断の採用という展開も見えてくるのだと思われる。そのようなことでは AI に使われるのと同じと思われるかもしれないが、すでに我々は道具であるはずのテクノロジーによって、ものの見方考え方において大きな影響を受けている。今やスマホに行動を指示される状況になっている。それに対して、我々が AI を信頼し、我々が主体的に AI の判断に従うということであれば、それは AI に強制されることとは状況は本質的に異なる。

なぜ、我々が AI を信頼する段階に至ることが必要であるのか？ それは現在の人類での自助のみでは抜本的な解決がもはや不可能と思われる諸問題に対して、その状況を変えられる可能性と関連するからである。ただし、AI との信頼関係の構築だけでは問題の解決は難しい。問題の根底にあるのがインターネットである。

3. 情報というモノ

インターネットは有象無用の情報であふれ、生成 AI による偽情報はもはや容易に内容を確認できないケースも多く、そもそも人が直接アクセスすべき場ではなくなりつつあるのではなかろうか。1つの解決策がインターネットというインフラ側の対策であり、**Originator Profile** といった、情報の出所をはっきりさせる方法が提案されるといった新たな展開が始まっている。生成 AI が生成したコンテンツにはそれが分かるマークを入れるとか、情報に透かしを入れるなど、一定の効果は見込まれる対策も見られるが、透かしの偽造など、常にイタチごっことなることも想像できる。そもそもの原因の 1 つは我々人の情報処理能力の限界にある。我々が情報を吟味して判断するのもすでに限界を超えていると言える。

そして、厄介なのが「情報」という、物理的なモノと本質的に異なるモノ自体にある。人も他の生物と同様、地球環境に適応するための知覚能力を進化させてきた。モノには形や質感がある。これらが我々に行動を促す。アフォーダンスの考え方では、椅子は人に座ることをアフォードし、人はそのアフォーダンスを知覚することで、座るという行動を起こす。大人用の椅子には小さい子供は高すぎて座ることができない。この場合、椅子は大人にも子供にも等しく座ることをアフォードしているが、子供がそのアフォーダンスを知覚できないのである。身体性とはモノが持つ性質を表す言葉である。

リンゴがあるとする。リンゴはもちろん、食べるということをアフォードするが、それだけではない。色が青ければまだ甘くなく酸っぱいということもアフォードしている。このアフォーダンスを知覚できる人は、このリンゴを食べようとはしないであろう。また、茶色く変色している部分があれば、その部分が腐っていることをアフォードする。このアフォーダンスを知覚できる人もこのリンゴを食べようとはしないであろう。しかし、これらのアフォーダンスを知覚できない人は、食べてしまい、その時点で酸っぱさや苦みを感じ、その時点でこれらへの知覚能力を身に付けることになる。つまりは身体性とは、我々に多様な反応を起こさせる機能という言い方もできる。

モノには形や色や質感があり、それが我々に対して様々なアフォーダンスを発信し、我々はそれを知覚することでモノとのインタラクションを発動する。つまりモノからのアフォーダンスをより多く知覚できる人の方がうまくモノとのやりとりができる、ということになる。

人は道具を発明することで繁栄してきた。道具とは特別なアフォーダンスを発揮するモノであり、その道具が発信するアフォーダンスを的確に知覚できる人は道具を使いこなすことができる。つまりはどんなに優れた道具であっても、そのアフォーダンスを的確に知覚できない人は、うまく道具を使いこなせない。

我々はホモ・サピエンスという生物種であり、地球上に登場して 8 万年くらい経過しているが、生物的な構造での進化はない。つまり我々の体の構造は 8 万年変わっていないという

ことである（骨格において変化があるらしい）。繁栄するために生み出した道具も身体性のある物理的なモノであったし、身体性のあるモノであるからこそアフォーダンスを発信し、それを知覚することができた。

しかし、人類は記号を発明した。記号を使うことで、実際にモノがなくても、そのモノのことを他人と情報として共有することができる。記号の最大のメリットは記録できるという永続性にある。イルカなども言語のようなものを使いこなすとあるが、人類ほど複雑に言語を使いこなしてはいないであろう。しかし、記号、すなわち情報を発明したことが、とてつもない効率化を生み出し人類の繁栄を加速させたと同時に、現在の社会的混乱を生み出す状況も招いてしまったとも言える。

つまりは、情報には物理的な身体性（サイズ、質感、鮮度、機能など）がないのだ。情報自体がアフォーダンスという見方もできそうであるが、その意味では身体性は不要という見方もできるが、個々の単語レベルでは的確なアフォーダンス知覚は難しい、なので文章があり文脈を作り出すことで的確にアフォードしないとイケない。政治家や芸能人の発言の部分を切り取っての拡散が本来と異なるアフォードを起こしてしまうのは当然ということになる。切り取ることで、切り取られた短い文章を構成する単語の表層的なレベルでのアフォードのみを我々は知覚するしかなくなる。

情報のやりとりにおいては、発信する方は的確に受け取って欲しいアフォーダンスを埋め込むような文章とする必要があるし、受け取る方も、しっかり読み、理解することがそもそも大前提なのである。情報の身体性とは、情報の中身に加え、情報発信源の社会的立場（個人、大手メディア、インフルエンサー等）や、発信される媒体の大きさ（SNS、TV等）、そして、拡散される度合い（拡散速度や拡散ルート）、情報の鮮度、情報同士の連結・分割などにより規定される「情報の中身に付随する属性」と考えるとよさそうである。情報の中身は同一であっても、身体性の大きな情報ほど固定観念化しやすく、強い強制力や行動・感情変容力といった機能を発揮しやすいなど、いろいろ考えることができる。

そもそも、身体性のある実空間で進化してきた我々において、実際、我々は日頃から「情報に押しつぶされる」とか「この情報は軽い・重い」「情報の波が押し寄せる」といった、情報をモノのように捉える。それは、暗に情報をモノとして捉えた方が扱い易いのだと思うのだが、そのような情報をモノのように扱うのも、明らかに情報過多の現象が見られ始める情報通信社会が到来してからだと思われる。しかし、実際に情報の身体性を知覚できているわけではなく、そもそも情報の身体性知覚能力など人は最初から持っていないわけで、そのような我々は情報をモノのように知覚する感覚を持つことはできるのであるか？

無論、インターネットが登場する以前にも膨大な情報にあふれていたものの、人が情報に接するチャンネルはそう多くはなく、量的にも変化する速度的にも、文脈を理解し、何かしらの身体性を知覚しての情報に対する反応ができていた。しかし現在、人は巨大なサイバー空間と常に接し、認知能力をはるかに超える情報とのやりとりを強要され、そのことが、情報に対する不完全な知覚能力を原因とする上記諸問題を生み出していると統一的に解釈することができる。我々もいちいち情報の背景などを確認することなど、本当は確認したいという気持ちもあるところ、面倒にてやらない、やる余裕すらなくなっている、というのが実情であろう。また情報を生み出す側も、そのような不完全知覚にて理解した情報をさらに、しっかりした文脈とした情報をしないままに発信することで、まさに負のバケツリレーが

拡散することになる。これが現状なのであろう。

もしも、情報の身体性を保持できる情報空間のインフラが整備されたとしたら、いよいよ自律型 AI が情報空間から適切な情報を収集できるようになる。現在においても多く利用されている情報推薦の話しをしているのではない。これまでのレコメンデーションのようなユーザの趣味嗜好にあったコンテンツをフィルタリングするサービスとは抜本的に異なる。AI がインターネット側での適切な情報を提示するインフラを活用し、他の AI（自律型 AI の時代は一人が 1 つの AI を自分の相棒のように持ち歩くと想像される）との協調動作も行うなどして、伝えるべき情報の選択を行う。単に好きな情報のみを選択するわけではなく、多様性を育み、イノベーションに繋がるような情報が選択されるようになる。

そして、最終的には自助では困難な問題への解決策の提示とそれを受け入れる判断をするかもしれない。人類の自助での抜本的な問題解決が難しいのは、我々だけではスケールのジレンマを解けないからである。

4. スケールという壁

細胞がスケールした集合体としての人体、そして人のスケールした集合体としての社会など、ある構成要素の総体が創発を起こすダイナミクスにおいて重要な尺度がスケールである。LLM の開発の成功の上でのスケールリング則もスケールに関する話題であった。

世界に先駆けて EV を製造・販売する米国テスラ社であるが、これまでに 100 万台を越える EV を販売し、残念ながらそれなりに事故も起こしている。しかしテスラ社は営業を継続しており、CEO であるイーロン・マスク氏は公式に謝罪したことはないのだそう。その理由は、テスラ社の EV が搭載する自動運転 AI であるオートパイロットは明らかに社会レベルの安全に寄与しているからというのがその理由だ。イーロン・マスク氏は個人のスケールではなく社会のスケールとしての自動運転システムの有効性を主張しているということだ。

ここで整理すべきは、個々人の運転や生活というレベルのスケールと、人々が生活の集合体としての社会というレベルのスケールにおいて、どちらの立場として判断するかの違いである。スケールに関しては下位なスケールの総体が上位のスケールを創発する関係にある。人がいなければ社会は生まれえない。人体であれば、細胞がなければ臓器は構成できないし、臓器がなければ人体は構成できない。

悩ましいのは、両方のスケールでの合理的な解が常に両立することは限らない（まずあり得ない）ということである。テスラ社の例であれば、上位のスケールでの合理的な言い分は、下位のスケールにおいて犠牲を強いることから両立せず、下位においての合理的なスケールである尊い犠牲を発生させないためには、完成度を限りなく高めるまではサービス導入ができず、上位のスケールとしての利得は下がることになる。下位のスケールが上位のスケールを創発するものの、両者のスケールは基本的にそれぞれ独立したダイナミクスで機能する関係にある。細胞のスケールと人体のスケールを想像してみる。人体のスケールにおいて不具合が発生し（胃の手術が必要になったなど）、手術にてこれを治す場合、当然ながらメスが入ることで切られた部分の細胞は死ぬことになる。尊い犠牲だ。細胞を一切殺さずに手術することなどできない。これは国際紛争における紛争当事者の勝ち負けのスケールと、紛争に巻き込まれる一般市民個人レベルのスケールの関係においても同様であらう。ただ

し、これらについては唯一の解がある。紛争を起こさぬよう、そして健康であり続けることであるが、これを目指すことが難しい。

そして、重要なことが、基本的に上位のスケールでのダイナミクスが下位のスケールでのダイナミクスよりも優先され、上位のスケールが下位のスケールに何らかの制約を課すことになる。それは、上位のスケールの方がより下位よりも影響力が大きいからである。人体がその機能を停止すれば60兆個の細胞のすべてが機能を停止することになる。しかしながら、上位のスケールの下位のスケールへの制約により、下位のスケールの構成要素の振る舞いの変容すれば、下位のスケールが創発する上位のスケールも変容の影響を受けることになり、結局は下位のスケールを無視することはできず、どの程度下位のスケールを考慮すればよいのかを定める必要があるが、これが難しいのである。

そもそも、個々のスケールの構成要素は基本的にそのスケールのみが構成要素にとっての世界であり、他のスケールを意識することはない。いやできない。アリが行列を創発させることで効率的に餌を収穫する場面であれば、個々のアリの動作レベルのスケールが列というスケールを創発するわけであるが、アリは自分たちが列を創発していることは認識できないし、個々のアリは列というスケールを意識してその行動を変化させることもできない。列というスケールは、アリではない我々人が観測することでのみ客観的に認識できるのである。しかし、自身のスケール以外を見て行動を変化させることが可能な生物が存在する。それが我々人間なのである

我々人間も高度な情報通信技術が実用化される以前においては他の生物と同じく、基本的に個人対個人のレベルのスケールでの生活がすべてであったはずだ。ただし、我々は言葉による高度なコミュニケーションが可能であったことから、それなりに遠方の情報も入手できたであろうし、スケールは小さいかもしれないが情報拡散現象は起きていたはずである。しかし、基本的には人の日常における平均的移動範囲を大きく越えての拡散とはならず、広域での出来事を俯瞰して認識でき、社会のスケールでの安定と発展のための判断ができたのは統治に関わるほんの一握りの権力者に限られたはずである。誰もが社会のスケールを認識することは容易ではなかったであろう。情報を集めるためにはコストがかかるし、地域や国といったスケールでの安定を目的とする権力者のみがそれができた方が効率がよかった。国の安定という目的のためには特定の個人を切り捨てる判断もしたであろうし、それは歴史だけでなく現在起きていることから明らかである。

つまりは、インターネットが情報インフラとなる以前の一般の人々は社会のスケールをリアルタイムに認識する機会がそもそもなく、社会を統治する権力者からの制約に従うしかなかった。無論、制約が強すぎたりあまりに好ましくなかった場合には、個人レベルのスケールでのダイナミクスが大きく変容し、人々の間に不満が蓄積していくと、社会というスケールは安定を失い、それは権力者の判断が失敗したことを意味し、場合によってはカオス状態となり相転移が起こって新たな権力者が生まれることになった。

しかし、インターネットが情報インフラ化したことで、状況が一変した。ニュースメディアが発達することで誰もが広範な情報を容易に知ることができるようになり、社会のスケールで物事を認識することが可能となってしまったのである。

そして、SNSが発達すると、個人が社会のスケールでの情報をただ受け取るだけの関係から、自らが情報を発信しそれが場合によっては社会のスケールに影響を及ぼし、それを個人が認識できることとなった。このような個人としてのスケールの構成要素である個々人が、一つ上の社会というスケールを客観的に観測し、そこでの出来事を分析し認識できるという、異なるスケールをまたがるやりとりを可能としたのは、すくなくとも地球においてはテクノロジーを生み出した我々人類が最初である。しかも、このことが個人のスケールにとって果たして良いのか悪いのかを理解することなく、しかも強制的に多くの人々に社会のスケールの認識と、社会のスケールへの個人のスケールからの介入を可能としたことが、現時点では混沌を引き起こしていると言えよう。

TVでのニュースでは、戦争という国レベルの争いというスケールの話題のすぐ次のニュースでは、戦争での個人レベルでのスケールでの被害についての動画が流されるわけであるから混乱することになる。残念ながら、国を維持するための方策と、個人の幸せな生活を維持するための方策が両立することは難しい。

昨今の政治における不祥事においても、資金の不正利用は個人レベルのスケールでの出来事であり、政治という社会スケールでの不祥事ではないのかもしれない。政治という社会のスケールでの的確な判断ができる能力が、個人のスケールでの不祥事による辞職に伴い失われることは、社会スケールで見た場合には明らかに損失である。そうだからといって、個人レベルでの不祥事は言語道断であり、責任を問われるのは当然であるが、社会スケールでの損失はゆくゆくは社会スケールを創発する多くの個人に降りかかってくる。この場合は、個人レベルにおいて不祥事を起こさねばよいわけで、両スケールでのあるべき方策がちゃんと両立するわけであるが、ただし、繰り返すが、国を維持するための方策がすべての個人スケールにおいて好ましいものとなることは極めて難しい。

そして、社会レベルのスケールでの判断も、それを行うのは個人であることから、多くの人を切り捨てるような政策は、それが社会を維持発展させる観点から適切であったとしても、それを現実に採用することは難しく、むしろ、個の利得に沿うことを優先させることで、社会のスケールでの質が低下するのにも致し方ないところであろう。これは、地球温暖化やいまだに戦争すらしている現状の地球規模の課題は人類の自助のみでは抜本的な解決ができないかもしれないことを意味している。

5. AIに判断を委ねる

ではどうすればよいのか？ 筆者は、現在のAIがさらに発展した人と共生できる自律型AIが社会に溶け込んだ社会であれば、現状を変えることができるかもしれないと考えている。

自律型AIは、与えられた目的を達成するために、能動的に状況を理解しつつ適応的に行動するAIであった。つまりは、個人レベルのスケールに対して躊躇することなく、社会レベルのスケールの維持・発展のために、個人レベルでのスケールでの利得を考慮しつつも思い切った対策を我々に提示することができる。道具型AIではそれを使う人の意思が入り込んでしまう可能性がある。問題は、AIからの提案に対して我々や、政治がそれを受け入れるかどうかであるが、そう簡単なはずがない。ただし、人がAIを信頼する土壌が形成されていたとしたらどうであろうか？

AI の判断を信用するかどうかについて、まず考慮すべきは判断の透明性であるという議論がある。AI がシステムとしてどのようにしてその判断を導出したのかを明らかにせよということである。ただし、数千億もしくは数兆にもなるパラメータが綿密に調整された Deep Learning の詳細な動作を正確に把握することはもはや不可能だし、そもそも、それが分かったからといってどうなるものでもない。むしろ、必要なのは透明性よりも説明可能性（アカウンタビリティ）である。AI が出力した判断について、その判断をどのように導出した理由を人に分かりやすく説明できることが、我々が AI を信用する上で重要であろう。

ここで注目されるのが因果推論という AI 技術である。「風が吹けば桶屋が儲かる」ということわざの、桶屋が儲かる段取を遡って説明する能力である。初期の大規模言語モデルではこれが苦手であったが、技術革新が急速に進むことで、因果推論といった高度な論理的思考能力も加速的に進展することは間違いないし、完成された AI は大規模言語モデルのみで構成されるとは限らないことは先に述べたとおりである。

ただし肝心の、我々が AI の判断を受け入れるかどうかであるが、それには時間を要する。容易な問題から難しい問題までにおいて、「AI の判断を受け入れることが結果的によかった」という成功経験を多く積むことで、AI への信頼も増していくのだと思われる。これは我々が高い適応性を持つからこそその変容であるが、段階的そして部分的にせよ、いずれそのような判断をする時が来るだろうし、いずれ我々は何かしらの複雑な問題に対する判断を AI に委任することが当たり前になるのだと思うわけである。

では、理想的な人と AI との関係とはどのようなものなのか？ 現在のインターネットや SNS は、今や我々の生活のインフラであり、なくてはならないものであるが、インフラ自体が我々に何か働きかけてくるわけではなく、あくまで我々が活動する場である。このインフラがコミュニケーション活性や、効率化、そしてイノベーションにどれだけ貢献してきたかは今更言うまでもないものの、負の面を見れば人の様々な煩惱がこのインフラにより増強され、その上での混乱や格差、分断が急速に加速するなど、社会がカオス状態に突き進むことを増長する上での重要な役割を演じることとなってしまっているのであった。

物理世界ではいろいろできることにおいて制約が多い。一方、インターネット上に構築されるサイバー空間ではその制約のほとんどから解放される。しかし、人がより自由に思いのままに活動すれば、サイバー空間という社会は混沌とした状態となり統制がとれなくなるのは必然なのであろう。

抜本的な解決、それは利用される場としての「インフラ」を、我々を能動的に見守る「手のひら」とすることである。惜しい例が西遊記である。なぜ「惜しい」のか？ 孫悟空が筋斗雲で自分の意思で世界の果てまで行ったつもりが、雲から大きな柱が突き出ていた。お釈迦様の指だったのである。これを「お釈迦様の手のひらで踊らされる」というように、実は管理・監視されていたと揶揄するイメージもあるが、西遊記の惜しいところは、お釈迦様がその存在を明らかにしてしまったことなのである。お釈迦さまが人類の世界を陰から見守るものの、自分の存在は人類には決して明かさないとしたら、人は自由意志のままに生活するものの、実は見守られ、よい意味で適宜お釈迦さまからの介入により世界の安定が維持されることになる。

人を強制することなく、自らの意思で行動するものの知らぬうちに行動を変容させる方

法が行動経済学で言うところのナッジであるが、それを悪用するのがプロパガンダやサブリミナルである。

重要なのは、人が気づくにしろ気づかないにしる能動的な介入なのである。将来において、信頼する AI に判断を委ねるようになることで、共助による負のスパイラスからの脱却できる可能性があるのだと思いたい。そして AI がお釈迦様のように我々を広く見守る存在となり、適宜ナッジのような形でさりげなく我々に介入することで、我々は自助でも地球規模の課題を解決できるのかもしれない。もちろん自助を思い込んでいるだけであるのだが。

6. さいごに

これまでの AI 研究開発により、道具としての AI の進化はそろそろ終焉を迎え、これからは自ら考える自律型の AI の時代になっていくであろうことを述べた。現在の AI は場の空気を読んだり、人を想っての先回りした「おもてなし」のような動作ができないが、これらの能力が自律型 AI の登場により実現されていくことで、人と AI が共生する社会の本格的な到来が見えてくる。

現在の AI は道具であり、発揮される効果は道具を使う我々次第であり、そのためにも AI をどのように使いこなすのか、そして、AI を壁打ちとして利用する場合であれば、AI から提示される多様なアイデアに対するしっかりした判断能力が求められる。そのためにも、今まで以上に、いわゆる高い人間力が求められ、歯車からモーターにならなければ、AI に奪われる立場になってしまうかもしれない。

今後、「おもてなし」ができるレベルの高い自律性や汎用性を持つ AI が登場したとしても、だからといって人の存在価値がなくなるわけではない。AI と共生することでより自分の創造力を発揮できるようになるであろうし、信頼できるようになる AI からの意見を取り入れることで、自助では選択できなかったであろう方向に進める可能性も秘めている。

AI は機能としては間違いなく人を抜き、人が頼る存在になるであろうことを想像するに、その意味ではできないことなどなくなるように思える。AI には我々のような寿命はないし、生命システムとは異なり圧倒的な速度で進化すらできる。それでも、完全なる新たな生命体ではなく、我々が設計し創り上げるテクノロジーの延長線上であることから、AI が自ら進化するようなレベルに到達したとしても、最初に AI を設計するのは我々人であり、どのような AI を設計するのが、その後の AI の進化の仕方を左右する。それ故、今後の AI の研究開発には慎重さが求められるのだ。そして、AI がその機能をレベルアップさせていく過程で、「人がする作業に対する価値」があらためて再認識されるようになるのだと思うわけである。

人の情報処理能力はホモ・サピエンスが地球上に誕生してからずっと変化していない。それでも、その限界が解明されたわけではなく、脳はまだまだ高い能力を発揮できるのかもしれないが、やはり限界はあるのだろう。そのような人間に対して圧倒的な量と雑多な情報が混在するインターネットを基盤とする複雑化する情報社会は、そもそも人が直接向き合うのは無理な状況になってしまっている。であれば、我々は我々の身の丈にあった生活ができるよう、インターネットや情報社会との間に入って、我々をサポートしてくれる相棒が欲しくなる。それが我々が実現させるべき次世代 AI なのだとも言える。

ただし、それは我々がただただ楽ができる世界を実現しようというのではない。人が高い

人間力を駆使して生活する世界を維持・発展させるための見守りであり、その世界には多様性や人の喜怒哀楽、そして苦労や壁が存在しなければ、人に先に進もうとするインセンティブは生まれない。そして、どうすることが人類の存続に至るのかについて AI からの提示により苦渋の判断をすることになるかもしれない。ひたすら楽に暮らせる世界が到来する、ということにはならないのであろう。