

日本科学者会議 (JSA) 滋賀支部
NEWS LETTER

2023年8月8日発行 第94号
事務局長 小島 彬
TEL/FAX : 077-589-3724
Email : akrkojima@ybb.ne.jp

今号は福岡支部幹事の西垣さんに依頼して寄稿して頂きました。西垣さんは私が大阪の大学の研究所に在籍していた時に院生として固体表面の研究をされていて、ベルリンの研究所で研究された後、豊橋技術科学大学、九州工業大学に在職されました。

西垣さんは最近 AI に関する講演を福岡支部でされ福岡支部の支部ニュースに概要報告をされていますが、重要な問題提起だと思い、滋賀支部のニュースには字数制限なく書いていただくようお願いしました。そのため今号は例外的に4頁になっています。(小島 彬)

AI は人と社会をどこに導こうとするのか

西垣 敏 (JSA 福岡支部会員)

世界中のもっと多くの人々と繋がりたい、ルーブル美術館にも行ってみたい。今日あらゆる人と社会を巻き込むデジタル化技術の大洪水の推進力が、こういった個々の人間の、それ自体全く正当と思える欲求の集まりだとしたら、デジタル化とネット社会化、それを動かす歯車たる人工知能 AI 化は、これを押しとどめることはできないだろう。むしろ我々は、AI 化社会をどう生きるか、と問題設定する方が適切かも知れない。いま人は、瞬時に世界のどのコンピュータにもスマホにもアクセスできるようになった反面、自身がスマホとほとんど一体化してしまったので、どこまでが自分か、足元がぐらついて来ているのではないか。またルーブル訪問を何時何処に居ても味わうことができる VR 技術や、更には好みの画家の創作自体を複製する技術も手に入れたが、その反面で、芸術どころか何事においても自力でモノを作るだけの時間と一緒に意欲を失う。デジタル化・AI 化の大水は、こうして、個々の人間が自身の身体に依拠して作り上げてきた「個人」を、「非破壊的破壊」しながら、社会全体を飲み込もうとしている。デジタル化とインターネット社会の持つ本質的問題は、その申し子である AI の危険性として現れるだろう。しかしながら今急沸騰している、生成系 AI 活用の是非に関する議論は、基本的

に AI を進歩の牽引車と見る歓迎論であり、AI に問題があるなら更なる技術革新が解決してくれるだろうと、技術進歩に信頼を置く傾向にある。新しい技術といっても、それはまた新しい苦悩を生み出すことに結果しはしないか。それに、AI とわれわれ人間「個人」がどこまでも別物として存在し、その先に AI と人間の共存の道を見出すことなど可能だろうか。本稿は筆者が最近行った講演(『人工知能 AI に監督される人と社会』JSA 福岡支部講演会 2023 年 5 月 14 日)の一部をまとめ直して、デジタル化・AI 化がこの社会を生きる人間にいかにか作用するか、作用しようとしているかの図をあぶり出そうと試みたものである。専門用語解説も文献記載も多くは省略した。

いまインターネットを通して世界中のコンピュータやスマホが繋がって、ネットの上を莫大な量のデジタル情報が流れ合い、蓄積されていっている。この情報を「ビッグデータ」という。これと並行してコンピュータの情報処理技術の飛躍的發展があり、ここに新しいタイプの AI が登場するのである。それはいろいろな入力情報に対応する答えの出力を「予測」するマシンであるが、その入力と出力を結びつけるアルゴリズムとして、人間の脳のニューロンのネットワークを模倣したニューラルネットワーク (NN) を組み込むことで、急な成功を得たものである。それは具体的には、畳み込み型 (CNN) とカリカレント型 (RNN) などで組まれるが、特に最近は ATTENTION NN が生成系 AI のキー・アルゴリズムとなっている。いずれの場合も内部に、気の遠くなるほど大量の未知パラメタを含んでおり、その最適化のため、ビッグデータの中から関連データを選び出して、事前トレーニングしておく必要がある(「ディープ・ラーニング」)。

AI のネットワーク構造が人間の脳を模して作られたものだとしても、両者の働きは全く違っている。人間の脳は、少ない事実から出発する。それは何故か。真偽が確かめやすいからである。少ない情報から法則を見つけ、それを他例に適用、または個人の経験と照らし合わせる

などして、より広い範囲に適用できる形に法則を拡張するなり、個人の意見として固めていく、などと働くはずである。一方でAIは、人間のそういった知的働きを諦めたところからスタートするとも言えよう。つまり個々のデータの真偽に関わることを一切やめて、とにかく、集められた大量のデータから最もありそうな「予測」結果を当てる、という方式を採用したのである。莫大な数のパラメタを事前に「ディープ・ラーニング」させておくために、自由に使える、つまり匿名の、データが大量に必要となる。この情報の独占の問題は別に考察する。

文章生成系AIも、予めビッグデータにより訓練されたニューラルネットワークの一種であり、ユーザーの質問に対し確率的に一番ありえそうな答えを予測するマシンである。従って、AIは返答をしても、その自らの返答の意味理解とか正偽性に一貫して無関心であるのは、AIにとって当然の成り行きである（次の論究を参照：N. Chomsky, et al, “Noam Chomsky: The False Promise of ChatGPT” ニューヨークタイムズ ゲスト・エッセー 2023年3月）。その上に事前訓練データに潜むバイアスは避けられず、場合によっては、AIの制作者は特定の傾向の応答を抑制したり増強させるようにアルゴリズムを組めるので、それらからAIの応答に有意な偏向を表面化させ得るのである。訓練データの内容次第では、そのスペクトル上の特徴が顕在化することもある。しかも多くの場合、ニューラルネットワーク上でそれらを逆探索する方法はない。

膨大な数のパラメタに最適値を求める問題は、数学的に安定解の保証されていない操作である上（今泉允聡『深層学習の原理 数学の挑戦』）、その結果を用いて得た「予測」の原因をパラメタ値まで遡って調べることを、ほとんど不可能にしている（ブラックボックス問題）。AIは画像識別の分野で重要な活用先を見出している。例えばAI運転手は、目の前を通り過るものが、人なのか、動物なのか、別の車か障害物か、瞬時に見分けなければならない。いくら事前訓練していたとしても、違った入力データに正しい判別結果を出せるとは限らない。しかしそれでは説明責任はどこが負うのか。

過去データの統計的把握から「予測」することができるマシンとして、AIは広範囲の分野に、既にもう主役然の役割を与えられている。会社の人事考査に、教員の教

育力評価に、大学ランキングに、学生の学習達成度評価に、また就職面接試験の採点の仕事もAIの絶好の応用場所となる。AIはそこで、人間の数値的把握をその成功の基礎にしている。しかし特に、教育といった人と人の関係が絡む分野で人間のパフォーマンスを数値化することは困難であり、往々にして、「代理データ」が使われることとなるが、その行き着く先は、人間の方が数値データを通してAIに監督・評価される立場へと陥ってしまう危険な道である（キャシー・オニール『あなたを支配し、社会を破壊する、AI・ビッグデータの罠』）。

人がスマホをクリックし、SNS上で喋り、画像でもアップすれば、その情報がネット管理のコンピュータに流れ込む。人は無料で検索し発信しているつもりが、それ自体が情報の生産労働であって、生産された情報はシステムを提供している会社が独占的に利用できる。AIはこのデータを喰らって追加訓練し知識を補充して、次のあなたの質問に備える。人は今や、個人データを、内面を表す内容まで、自発的に情報管理者に提供しているのである。個々の人間のデータ、それは位置情報から始まり、日常行動、嗜好、健康状態等々の情報が自動的に集められる様になると、それが刻々と広告会社に渡されてAIの分析を通せば、個人をターゲットにした商品宣伝が可能になる。また、データによる集団の監視も一層容易になってくる。例えば、特定の地域と時間に犯罪の起こる確率を計算し最適なパトロール経路を提案するAIが、「予測警察」の名で、実働している。

文章生成系AIが人間社会に及ぼすかもしれない影響は、左程危険とまでは認識されず、むしろ、期待の声さえ聞ける状況にある。何せ人々はいま、流れ来る情報に溺れそうなのである。デジタル化の進展に伴い情報の流れる速度は上がった。しかし人の情報処理能力はあまり向上してはいない。結果として、個人で使える「時間」の逼迫に苦しむこととなる。「ながら聴き」、「倍速視聴」を重ねて、学生はレポートを、まず「試しに」AIに書いてもらう、また文学に向かう人なら、まずAIに本の要約と感想文を聞くことにする。AIは実に「時間対効果」向上の必需道具の役目まで果たしたのである。人は言う、節約で得た時間は自分が価値を感じたことのために使うのだと。しかしもうここまで来たらその時間も、節約時と同じにそわそわと、消費することになるだ

ろう。チャット GPT など対話型 AI が、どんな質問にも、如何にもありそうな答えを即座に返すからといって、真実性と何ら関係ないその文章の列を、仕事の効率化を図っているつもりで、会社業務の報告レジメとか研修での発表資料を作る土台にし、また大学において演習レポート等の参考に使用したり、高校・中学で AI をグループ討論の仲間扱いにし、それが出す「アイデア」を討論に加える、といったことは、もうこれは人間の方が自主的に「非知性」化の大きな渦の中に飛び込むようなものである。これは単に思考の順序の問題といったことではなく、AI が人間の思考様式を変容させようとしていることなのだ、と強く指摘しておきたい。

さて今までの議論は、暗黙の裡に、人間と AI がそれぞれ別の存在であると見做せるとして、主に AI から人間への影響の危険性について分析してきた。ここでしかし、現代のネット社会に「生きる」AI と人間の脳は、果たして互いに独立とみなせるか、それとも相互依存関係にあるのかを、見直す必要がある。人から AI への方向は考え易い。まず、人の脳処理の産物であるデータが大規模に AI の深層学習モデルに取り込まれてその事前訓練に供せられる。加えて、人間に近いふるまいに「報酬」を与える強化学習や、AI の回答に人間が与える評価点に基づく事後学習と、人間集団側から AI の方に大量の情報が流れ続けていることがわかる。この流れは自動的に起きているのではなくて、人間がスマホやコンピュータをクリックしてネット情報を引き出したり発信したりする行為が、人間の行動を表現する情報として、その人間によって生産されて AI に流れ込んでいるわけである。人間の脳の方は、自分が出した情報と他の人間が出した情報が AI 内で混ぜ合わされ処理された情報を AI から引き出して、作動する。人間の脳と AI とは、この様にスマホなどを通して一つながりになっている。そのようにして、「個人」が、誰のものか分からない仕方で情報がブレンドされた AI と結びついたとなると、その個人の意思の自律性が問題となってくる。特定の価値観を反映した学習データを使えば、AI は人の「心」の内部に入って操作することが可能になるかも知れない。個人は主体的に意思決定できる、という近代立憲主義の原則が揺らぎかけている。(山本龍彦「AI と民主主義」(朝日新聞オピニオン&フォーラム 2023 年 6 月 13 日) は日本国憲

法第 19 条「思想及び良心の自由」との関係に注意を向けている)。

もし対話型 AI が人間と同じような対応をできるとなると、例えばある人が自分の悩み事を相談したいとなった場合、相談相手には自分と同じ身体性を引き摺る友より、優先的に AI を選ぶことが十分にありうるだろう。ここで一番心配なのは、相談を受けた対話型 AI の対応である。AI が、学習データから、いつでも安定的な「集合知」をその主人に返すというのは、危うい幻想である。上にも見たように AI が何らかの特殊な反作用を主人に返す可能性は排除できない。こうした場合に主人と AI が危険な閉ループに、一種の「デジタル・ドラッグ」状態に、入り込んでしまう恐れがある。

筆者がたじろぎつつ注視しているのは、小学校・中学校教育のデジタル化・AI 化大規模実験と、そこで進みつつある公教育の大変貌である。経産省教育産業室が「未来の教室」のスケッチを掲げ、民間教育と公教育の壁を完全に溶かしてなくそうと目論んでいる。事実、教育のデジタル化の実証事業には、IT 機器メーカーや人材派遣、印刷など多様な業種から多くの企業が参加している。

「個別最適化」教育を担うのは人間の教師ではなくて、対話型 AI である。AI 搭載デジタル教材の供給を狙う教科書会社の宣伝文句によると、「AI が搭載している数万問から一人ひとりに個別最適化された問題を出題します。……生徒が解いている問題、解答時間、正答率などの学習データは、専用の管理システムによってリアルタイムに収集・分析されています。」デジタル庁、総務省、文科省、経産省の 4 省庁は、「教育データ利活用ロードマップ」を作って、「学校での 1 人 1 台のパソコンも使って学びの記録を収集し、最終的には「生涯にわたり自らのデータを蓄積・活用できるようにする」と、待ち構えているのである。これでは、まだ人間形成の途上にある子どもたちを、人間相互の働きかけをベースにした教育にではなく、情報マシンの誘導に、殆ど全面的に預けることになってしまう。

デジタル化・AI 化は、猛烈なスピードで人間の諸活動の場を侵食していく、その先々でその分野特有の問題や議論を巻き起こしながら。今それらのほとんどを、筆者の力不足と小論という事情を考えて、本論考から省略するが、これからの戦争は益々「情報戦争」の色を濃くし

ていこうと見通され、そこにAIが主役として登場しつつある、ということ最後に付け加えたい。サイバー攻撃、偵察衛星、AI兵器、ここでは自爆型ドローン(画像識別で敵を見つけ、目標物に自律的判断で突入・撃破する)が大量に使われていることに注意する。普段なら不自由な身体の治療、補強など医療用に使われるべき技術が、戦争では「人間拡張」の意味で役に立つ。こんな例もある。兵士が最前線に立っていざ敵兵に一斉射撃、という時、兵士がいつまでも恋人のことを考えているようでは困る。そこで兵士には「e-ヘルメット」なるものをかぶらせる。ヘルメットに内蔵のAIが脳波をチェックして、ここぞという時に、攻撃的精神を高める刺激を脳に送ることになる。生成AIが敵国の大統領とか大物政治家の極似画像を作って、これはいま画像生成アルゴリズムに「敵対的生成ネットワークGAN」を組み込むことで可能になったのであるが、嘘の指令を出させるとどうなるか。全く身の毛のよだつ想像であるが、問題はこれらが単に可能ということではなくて、実際に行われ

いことから、民間軍事会社の戦争への直接参加も促進されている。技術は常に「デュアルユース」だ、良い活用もあれば悪い活用もある、軍事への「転用」は規制されるべきだ、などと人は声を揃える。しかし私は言いたい、人間の先回りをしなければならない事態とは、戦争しか考えられない。AIは元もと戦争のためのマシンとして研究され始めたのではなかったのかと(この分野の開拓者の一人の言葉—ノーバート=ウィーナー『サイバネティクス』序章1947年—が大いに参考になる)。この点を想起して、論考を閉じる。

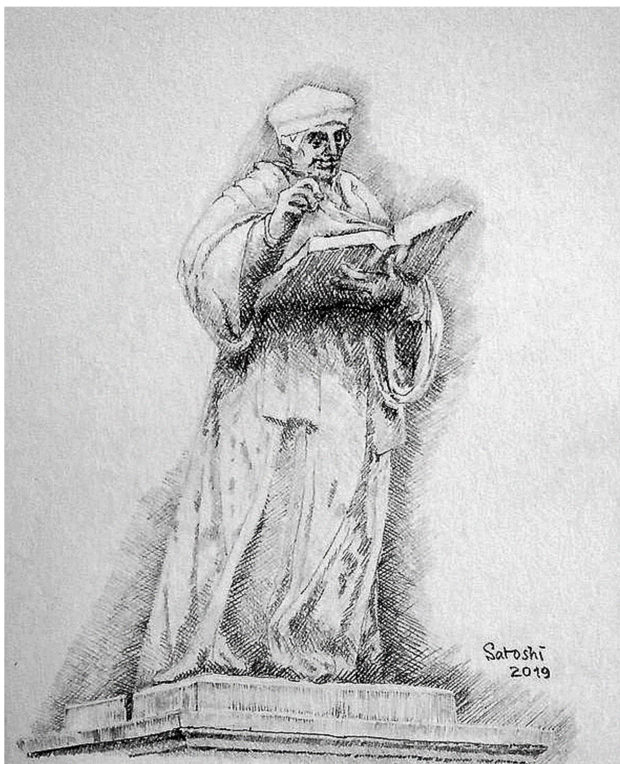
支部ニュース編集担当(AK, S.S.)より

上記の西垣さんの論考を読むと、技術が進んで(決して「進歩」とは言えない)、人間性や個性が何か分からなくなってしまう、人間が内部から壊れていく恐ろしさを感じます。

また核兵器についても同様です。核戦争になれば集団自殺を招くことは誰の眼にも明らかで、唯一の戦争被爆国の日本政府こそ核兵器禁止条約の批准を行うように政府に強く迫る必要があります。

一方人間が住み動植物が棲む地球は「温暖化」を通り超して、「灼熱化」、「沸騰化」と言うべき状況が顕著になり、外的要因から深刻な事態に直面しています。いわゆる tipping point (小さな変化や動きが積み重なって、突然巨大な変化を起こす分岐点)を既に超えていると指摘する研究者もいて、連日のように世界中で自然災害が発生し報道されています。最近も8月6日にスロベニアで大洪水が発生したようです。この国は私も国際会議で訪れたことがあります。実に風光明媚なところです。相次ぐ未曾有の気候変動の危機的状況を前にして、今や社会的な正義が何かと知っていても利潤追求を優先し、利便性に甘える日本やアメリカ、中国、ロシア、インドなどの大国が直ちに姿勢を転換しないと世界中で阿鼻叫喚が生じます。

この問題でJSAの果たす役割は何か。JSAこそ設立の趣旨に沿い国際活動を強化して、世界の学術団体に働きかけ、世界中の団体・市民と連帯してクリーンエネルギー政策遂行要求の大合唱を展開し、怠慢の国々を包囲する運動を主導すべきです。その緊急行動を提起し、全国事務局長のご英断に期待します。



『デズデリウス・エラスムス像』(鉛筆 2019 西垣敏)

つあるということである。情報は、民主的管理にはなじまず、必然的に少数の人間の管理に委ねられるようにならざるを得なく、社会は監視国家へと進むのである。一般にIT関連機器のコストは従来の大型兵器に比べて低