

生成Aーによる 人間は何を失うのか

酒井邦嘉（言語脳科学者）

生成Aーは、人類の夢をかなえる万能の技術だろうか？

言語脳科学者の酒井邦嘉は、その使用は言語能力だけでなく、

思考能力や独創性をも減退させると警鐘を鳴らす。

従来のAーに比べ、生成Aーがはるかに危険な技術であるのはなぜなのか。このままでは二〇二三年が人類の文明退化の分岐点となると警告する

酒井に話を聞いた。

——生成Aー、とりわけChatGPTが驚くべきスピ

ードで社会に広まりつつあります。酒井先生は言

語脳科学者であり、教育にも携わっていますが、生成Aーをどのように使っていけばいいとお考えですか。

現状のチャットボットの技術はリスクが大きく、特に教育全般で使うべきではないと考えています。

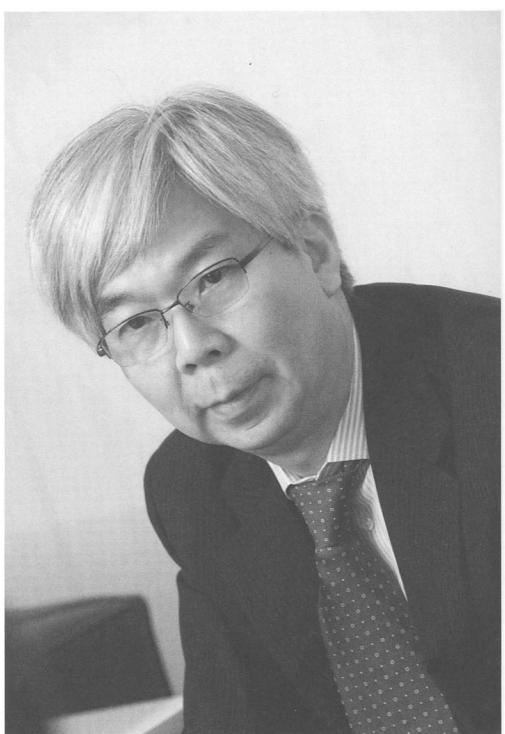
——それはなぜでしょう？ 生成Aーが提示する

回答に誤りが多いからですか。

それもありますが、対話風に仕立てられてい

ることで、人が「正解」を求めるように誘導されるリスクが大きいからです。「適切な問い合わせをすれば、ChatGPTはうまく答えてくれるだろう」と期待した時点で、すでにチャットボットの呪縛から逃れられなくなります。そもそも

ChatGPTとのやりとりは「対話」と呼べるものではありません。ChatGPTの回答を「託宣」



写真=幸田 森

に頼つて自分の頭で考えなくなる」のは必定でしよう。今ここで踏みとどまつて規制をしないと、もう後戻りができなくなります。

生成AIの限界

——AI技術はこれまでも活用されてきました。

従来のAIと生成AIは何が違うのでしょうか。

従来のAIは、入力されたデータの特徴や傾向を学習して、類似する別のデータを識別したり、先を予測したりすることが主な用途でした。

ChatGPTなどの生成AIは、そうしたAI技術の延長線上にあり、インターネット上の膨大なデータから条件にかなつたものを選んで文章や画像などを合成します。先読みの技術で読みやすい文が作れるようになつたとはいえ、文の構造や意味を把握しているわけではありません。多くの人が誤解しているようですが、AIや機械が「考える」とことなど決してありません。新しい組み合わせの文章や画像を合成できるからといって、それだけでは革新的な技術と言えず、そもそも真の「生成」ではないのです。

——新しい組み合わせを生むなら、創造的なのでありますか？

アメリカの言語学者ノーム・チョムスキーは、

「英語話者が新たな発話を産み出したり理解したり出来る一方、他の新たな列〔筆者注：音素や文字の列〕を英語には属さないものとして退けることが出来るという能力」が言語の本質だと述べています（『統辞構造論』ノーム・チョムスキ著、福井直樹・辻子美保子訳、岩波文庫）。

この「英語」は、あらゆる自然言語に置き換えて考えることができます。これが眞の生成です。

ChatGPTには、「新たな列を英語には属しないものとして退けることが出来るという能力」があります。構造や意味、そして論理の間違つた文章をいくらでも合成してしまいます。ですから、人間のような言語能力や創造性はない

と断言できます。創造性には、不要なものを捨て去る能力も必要なのです。

生成AIの文章は、もつともらしく見せかけた文字列にすぎません。いくらやりとりを重ねたところで、それは実際の人間同士の対話や、人間が生み出す創作物とはほど遠いものです。例えば、「彼のように失敗してはならない」と

いう文を考えてみます。はたして彼は失敗したのか、失敗していないのか、文字列だけではまったく判定できないでしょう。

これは、否定の範囲という構造によつて意味が変わる例です。「失敗すること」を否定するな

ら、「(失敗しなかつた) 彼のように、『失敗すること』はあつてはならない」という意味になります。一方、「彼のように失敗すること」を否定するなら、「彼のように失敗すること」はあつてはならない」となるわけです。

こうした曖昧な言葉はいくらでもありますから、文脈で判断できなければ話者に真意を確認する必要があります。論理学的に分析するには、「記述理論」による言語分析が必要です。しかし生成AIでは、文法的解析から意味の解析まで必要な分析が搭載されておらず、そうした文章は放置されます。

——そんなに不完全な技術であるのにチャットボットとの会話が成立するのはなぜでしょうか。

擬似的な対話風のやりとりだけがAIに実装されていても、その足りない部分は人間側がすべて補つて都合よく解釈しています。会話で相づちを打つように、言つたことを単にオウム返しするだけでも、ある程度は対話しているように見せかけられます。

しかし会話で最も必要なのは、相手の思考に対する推理です。相手がどういう意図で質問をしたのか、本当に聞きたいことは何なのか。ChatGPTはそうしたこと推測するわけではありません。

外国语について「日常会話くらいはできるようになりたい」と多くの人が言いますが、日常会話ほど難しいものはないのです。実際の会話では常に話題が変わりますし、ひとつひとつの言葉の意味も、文脈や解釈次第でさまざまに変化します。話し相手によって背景知識や発語傾向も大きく異なります。

文章作成や対話に必須な、構文解析・意図理解・解釈・表現といった言語・思考能力の再現がすべて先送りされていることからも、現時点での生成AIの限界は明らかです。

生成AIはカணニンぐやドーピングと同じ

— ChatGPTにはリスクがあるということですが、推進派の意見に対してはいかがですか。

そうした意見に共通して見られるのは、「新技術を恐れるな」「禁止しても何も始まらない」「上手に使いこなした方がよい」といった常套句です。

ChatGPTはマインドコントロールや洗脳にも悪用される恐れがあり、プロパガンダに使われれば、核兵器やロボット兵器と同様に世界を震撼させるでしょう。それでも「新技術を恐れる

な」「禁止しても何も始まらない」と言えるのでしょうか。

— 文部科学省からは今年七月四日に「初等中等教育段階における生成AIの利用に関する暫定的なガイドライン」が公表されました。慎重な対応を求めながらも、AI時代に必要な資質・能力の向上を図る必要があるとしています。

個人や家庭、そして学校や組織には、どんな道具や技術に対しても使用の拒否権があります。

リスクのある技術の積極的な活用を促したり、その活用法を声高に喧伝するのは、明らかに行き過ぎた行為です。

例えば、そのガイドラインにある「教師が生成AIが生成する誤りを含む回答を教材として使用し、その性質や限界等を生徒に気付かせること」という「活用法」は、それまでも繰り返し提案されてきたものですが、使用を前提として無理に提案された感が否めません。間違い探しはインターネット上の記事でも十分にできましすし、人が作成した文章の方がはるかに教育的でしょう。

教育上はつきりすべきことは、生成AIによる文章作成が「自らの実力を欺く不正行為」だということです。カணニンぐやドーピングと同じように、詐称の効果と、相応の副作用をもつ

【初等中等教育段階における生成AIの利用に関する暫定的なガイドライン】による、生成AIの活用が「適切でないと考えられる例」と「活用が考えられる例】

1. 適切でないと考えられる例

- ※ あくまで例示であり、個別具体に照らして判断する必要がある
- 生成AI自体の性質やメリット・デメリットに関する学習を十分に行っていないなど、情報モラルを含む情報活用能力が十分育成されていない段階において、自由に使わせること
 - 各種コクーンの作品やレポート・小論文などについて、生成AIによる生成物をそのまま自己の成果物として応募・提出すること（コクーンへの応募を推奨する場合は応募要項等を踏まえた十分な指導が必要）
 - 詩や俳句の創作、音楽・美術等の表現・鑑賞など子供の感性や独創性を發揮させたい場面、初発の感想を求める場面などで最初から安易に使わせること
 - テーマに基づき調べる場面などで、教科書等の質の担保された教材を用いる前に安易に使わせること
 - 教師が正確な知識に基づきメント・評価すべき場面で、教師の代わりに生成AIから生徒に対し回答させること
 - 定期考査や小テストなどで子供達に使わせること（学習の進捗や成果を把握・評価するという目的に合致しない。CBTで行う場合も、フィルタリング等により、生成AIが使用する状態とならないよう十分注意すべき）
 - 児童生徒の学習評価を、教師がAIからの出力のみをもって行うこと
 - 教師が専門性を發揮し、人間的な触れ合いの中で行うべき教育指導を実施せずに、安易に生成AIに相談させること

2. 活用が考えられる例

- ※ あくまで例示であり、個別具体に照らして判断する必要がある
- 情報モラル教育の一環として、教師が生成AIが生成する誤りを含む回答を教材として使用し、その性質や限界等を生徒に気付かせること。
 - 生成AIをめぐる社会的議論について生徒自身が主体的に考え、議論する過程で、その素材として活用されること
 - グループの考え方をまとめたり、アイデアを出す活動の途中段階で、生徒同士で一定の議論やまとめをした上で、足りない視点を見つけて議論を深める目的で活用されること
 - 英会話の相手として活用したり、自然な英語表現への改善や一人一人の興味関心に応じた単語リストや例文リストの作成に活用されること、外国人児童生徒等の日本語学習のために活用されること
 - 生成AIの活用方法を学ぶ目的で、自ら作った文章を生成AIに修正させたものを「たたき台」として、自分なりに何度も推敲して、より良い文章として修正した過程・結果をワープロソフトの校閲機能を使って提出させること
 - 発展的な学習として、生成AIを用いた高度なプログラミングを行わせること
 - 生成AIを活用した問題発見・課題解決能力を積極的に評価する観点からパフォーマンステストを行うこと

ているのです。使用を繰り返すことによる思考力の低下も明らかでしよう。それを初等教育から推進しようというのですから、常軌を逸しています。

なぜそこまでして害のある技術の利用を「暗

に使いながら、「それも校正の範囲です」とうそふかれたら、どうするのでしょうか。つまり、少しでも生成AIの使用を許容すれば、「何%までは使ってよい」という線引きができない以上、全

将棋A-1とチャットボットは別物

——生成AIの利用を推奨する意見の中には、将棋のAIを引き合いに出して、「プロ棋士が使って

いるんだから、生成AI一だつて同じように使えるね」というものがあります。

それは明らかに誤解です。囲碁や将棋などの

ケーブルA.I.はチャットボットとはまったく別物なのです。ボードゲームには勝ち負けという

明白な評価がありますから、現局面から何億手

という指し手を調べることで、有力な手をいくつ見つけられるか。

つか見つけることができます。

うした評価ができません。作品の価値判断にも

解釈や好みの余地がありますから、不動の「〇

○点] ということは難しいですし、そもそも厳密な数値化などできな、相談です。

習ふ數値化されない棋譜

対局で使えばカンニングと見なされることです

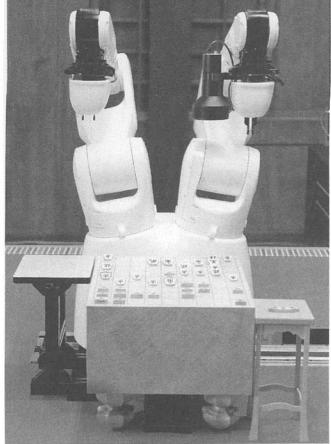
それなのに「使えばよい」という方向に誘導するのよ、多分の事例の「」から目脱二部合ひ、

るのには多くの事例の中から自説に都合のいい事例のみを提示する「チエリーピッキング(cherry picking)

——青少年読書感想文全国コンクールを主催する全国学校図書館協議会の談話として、「自ら書いた文章の推敲過程で、校正に生成AIを使うことまでは否定しない。これまでも保護者や教員が添削して作品を提出してきた例はあり、それと大きな差はないのでは」との報道記事（二〇二三年七月二日、朝日新聞朝刊）がありました。

「校正」といっても、ほとんど書き直しに近い推敲までが含まれます。生成AIの対案を大幅

ない人は、はたして教育者なのでしょうか。



将棋AI「Ponanza（ポナンザ）」は、2017年の第2期電王戦二番勝負で現役名人（当時）の佐藤天彦九段を2戦全勝で破った。明確な評価関数によって形勢を数値化できるからこそその成果。AI側は指し手もロボットアーム「電王手一二さん」が務めた。写真提供：毎日新聞社

先ほどの文部科学省による「暫定的なガイドライン」では、「活用が考えられる例」として、さらに奇妙な可能性が挙げられていました。「英会話の相手として活用したり、より自然な英語表現への改善や一人一人の興味関心に応じた単語リストや例文リストの作成に活用させること、外国人児童生徒等の日本語学習のために活用させること」です。ChatGPTが語学に使えるなどというのは、人間の言語能力や外国語を軽んじた誤謬です。

繰り返し述べてきたように、ChatGPTには本質的な言語能力が搭載されていませんから、たとえ翻訳や要約リストの作成であっても、単に「それらしいもの」でしかありません。

――企業や自治体が生成AIを導入するのは、何か問題がありますか？

――企業や自治体が生成AIを導入するのは、何とも思っていません。ただ、生成AIではうまく対応できなかつたとき、それを的確に判断して担当者につなぐことができるかは大いに疑問です。医療相談や教育相談などの窓口でChatGPTが使用されたら、頭部外傷やいじめといった重要な案件が放置され取り戻しがつかない事態を招くかもしれません。そうしたときに、「ChatGPTが初期対応を誤ったためです」などと、誤解は誰にもできないのです。

企業のカスタマーセンターや、自治体の相談サービスでは、窓口業務の人手不足解消や効率化を図つて、メールやSNSを通しての問い合わせに対し生成AIが次々と導入されていくと予想されます。これは、自動音声による対応よりもはよいかもしませんが、客や市民に対して「人として」の対応を放棄していることに変わりはありません。

もし生成AIではうまく対応できなかつたとき、それを的確に判断して担当者につなぐことができるかは大いに疑問です。医療相談や教育相談などの窓口でChatGPTが使用されたら、頭部外傷やいじめといった重要な案件が放置され取り戻しがつかない事態を招くかもしれません。そうしたときに、「ChatGPTが初期対応を誤ったためです」などと、誤解は誰にもできないのです。

ChatGPTの潜在的な危険

――私の知人はChatGPTを相手に、仕事の愚痴をこぼしているそうです。「今日は上司にこんなことを言われた」と書くと、「それはひどい。あなたの上司は間違っています」などと返ってきて、それがストレス解消や癒やしになつていているそうです。

ChatGPTは「イエスマン」として振る舞うようになりやすいのですが、それが手放せなくなつて依存してしまうと、もはや家族や同僚からの時には厳しい意見に対しても一切耳を貸さなくなりってしまう恐れがあります。「ChatGPTだけが私を理解してくれている」と錯覚して妄想が膨らんでいったら、カウンセラーの介入も受け付けなくなるでしょう。ChatGPTを搭載した悩み相談のソフトは、両刃の剣です。

日本では当時ほとんど報道がありませんでしたが、今年の三月下旬に、ベルギーの三〇代男性が、AIのチャットボット「チャイ」と六週間にわたつてやりとりを続けたあと、自殺するという事件がありました。その男性は当初、気候変動について問い合わせをしていたのですが、やがて「このままでは人類はどうなつてしまふのか」という不安が増し、架空の女性「イライザ」に相談するうちに一方的に感情移入をしてしまい、自死を促されてしまったそうです。

この事件はEUの欧州議会でも取り上げられ、生成AIの規制を強化する一因になりました。これも生成AIの潜在的な危険を物語っています。常に同意を返してくるだけの標準的な仕様であつても、当人が抱く不安定な感情を增幅さ

せることになりかねません。ChatGPTの利用を推奨したり導入を推進したりする人は、そうした事態の責任を取れるのでしょうか。

——生成AIが犯罪に利用される危険性も指摘されています。

音声や画像、そして映像までが、生成AIによって巧妙に合成できるようになつてきました。

今や、ある人の肉声の一部を使って、本人が全く別のことと語っているように音声を合成する技術があります。スタントマンなどの代役立てて、俳優の顔に上げ替えることもあります。そ

うした加工や合成を簡単に行うアプリが出回れば、誰でも「にわか監督」になれそうです。

すると、声紋データや証拠写真、そして防犯カメラの映像までが巧妙にねつ造される可能性が出てきます。そうしたものがひとたび信憑性をもてば、新たな冤罪(えんざい)が生まれるでしょう。その一方で、デジタルデータが犯罪の証拠としての価値を失う可能性もあります。

これまでさまざまなかつたわけですが、今までとは桁違いの詐欺や隠蔽(いんぺい)が横行し、それが犯罪組織によって大規模に利用されるかもしれません。

——先端技術に対して、なぜそうした負の利用が起ころのでしょうか。

先端技術の軍事応用は、すべて例外なく負の利用です。海外では銃の乱射事件があるたびに銃規制について議論がありますが、銃の保有が犯罪の抑止力になるという意見も絶えません。この点は核兵器やロボット兵器も同様ですが、抑止力になるからなどと言つて、全体的な危険を顧みないのは明らかな間違いです。毒ガス兵器から化学兵器、生物兵器まで、こうした技術を禁止する以外に人類の未来はありません。

かえりみないのは明らかな間違いです。機械に頼るようになつてしまつた代償として、若いイヌイットたちは、昔から受け継がれてきたサバイバル能力を失つてしまつたわけですね。

——一見、便利な技術でも、むやみに広めてはいけないのでしょう。

「イヌイットの遭難」が教えるもの

また、技術の進歩が人間の能力を衰えさせるという事例もあります。極北の地に暮らすイヌイットは、伝統的に犬ぞりを使って狩りをしていました。彼らは星や太陽の位置から方角を把握し、雲や風向きから気象の変化を予測して、昔ながらの知恵を生かしながら生きてきたのです。ところが一九九〇年代に入つてから、欧米のセールスマントリルがスノーモービルやGPSを売り込み始めました。これらの先端技術は狩りの範囲を飛躍的に拡大させ、効率的な収入増に役立ちました。ところがほどなく、遭難事故が多発するようになつたのです。

たびスノーモービルやGPSが故障してしまえば、致命的な事故につながります。機械に頼るようになつてしまつた代償として、若いイヌイットたちは、昔から受け継がれてきたサバイバル能力を失つてしまつたわけです。

日本では小学校から算数の授業に導入されていいる「電卓」も同じことです。計算の仕方を教える一方で、電卓の「積極的な利用」を進めるのは、明らかな矛盾ではありませんか。児童に電卓を使わせることで、「時間をかけて計算するのは無駄なことだ」と教え込むことになります。同様に生成AIを使わせることで、「時間をかけて文章を書くのは無駄なことだ」と教えることになるでしょう。

ただ、数学では、概念の整理や定理の証明といった考え方を重視されますから、筆算が多少苦手でも目立たないのでしょう。しかし、文章を書く力がなければ考えたことを言語化できませんから、あらゆる教科に深刻な影響をもたらします。

「電卓を教室で積極的に利用するなら、生成AIも同様にすべきだ」といった議論がまかり通つてはいけません。ここで、「筆算」する力と

「文章」を書く力を短絡的に同一視するのは、明らかな誤りです。言語能力は、あらゆる学びの基礎にあるのです。

このように考えれば、人間が考えなくて済むための「生成AI」は不要だということになります。

生成AIによる文明退化

——「文は人なり」というように、文章には書く人の個性が表れます。ChatGPTを使った人の個性はどうなるでしょうか？

当然、個性やその人らしさは薄れるでしょう。すでにChatGPTの利用法講座まで現れていますが、私は違和感を覚えます。生成AIがワープロソフトに搭載されようとも、それは決してワープロの延長ではありません。そもそも、そうした「借り物」の文章を自分の創作の一部とすること自体、学術的には正当な引用ではないのです。「オリジナルの文章を書く」という意識をなくしてしまえば、剽窃や盗作に直結してしまいます。

インターネット上の匿名記事やレビューでは、「誰が書いた文章か」という責任の所在がすでにわからなくなっています。そうした風潮に生成

AIが拍車をかければ、文章 자체に価値がなくなりますし、書き手と読み手の信頼関係も地に落ちるでしょう。悪意ある中傷記事や、愉快犯によるデマが大量に「生成」され、それが興味本位の大衆によって連鎖反応的に拡散されるのです。

それでも信頼のおける発信や創作を守つていくには、どうしたらよいでしょう。新聞記事や

小説の最後に「生成AIは一切使用しておりません」という但書きをして、デジタル署名でもつけましょうか。

絵画や音楽において過去の芸術作品に対する改変は、それが芸術家の手によるものならば「オマージュ作品」となりえます。しかし生成AIによる加工は、オリジナルの作意を歪めた形で広めうる破壊行為ですから、法的な規制が必要です。そうした加工画像や音源が生成AIで安易に作られてインターネット上に氾濫してしまつたら、悪貨が良貨を駆逐するように、作品の価値や評価を著しく下げてしまいます。

——生成AIの技術が広まれば、今後、世界はどうなっていくのでしょうか。

おそらくこの二〇二三年は、生成AIによる文明退化が始まつた年として後世に記憶され、これまでの人間中心の時代との分岐点となるでしょう。

よう。生成AIに頼るあまり、自分の頭で考えること自体を拒否し面倒がる人が続出すれば、教育としては悲劇的な結果を招きます。そして、効率を追求し、コスパ（コストパフォーマンス）やタイプ（タイムパフォーマンス）を重視して生成AIを使つた代償として、人々の思考力や創造力の減退に拍車がかかりますから、社会としては破滅的です。

生成AIによつて人間が失うものはきわめて大きいのです。近い将来、新たな社会問題や環境問題、そしてパンデミックなどに直面したとき、それに打ち克つだけの知力は、もはや人類に残つていなかかもしれません。

多くの人がAIの可能性に対し樂観的に考へていると思います。しかし、ことここに至つては、大きな危機感をもたずにはいられません。私はひとりの科学者として、生成AIといった技術を妄信することがないようにと啓発する必要を感じています。

構成・文＝布川剛

K

さかいくによし 言語脳科学者、東京大学大学院教授。一九六四年、東京都生まれ。東京大学大学院理学系研究科博士課程修了。一九九六年、マサチューセッツ工科大学客員研究員を経て、二〇一二年より現職。脳機能イメージングなどの先端的手法を駆使し、人間にしかしない言語や創造的な能力の解明に取り組んでいる。著書に『チヨムスキーと言語脳科学』（インターナショナル新書）、『科学者という仕事』（中公新書）など。