

## 特集 生成系AIへの対応

# 進化を続けるAIと人間 どう向き合えばいいのか

国立情報学研究所  
情報社会相関研究系教授 佐藤 一郎



ChatGPTは2022年11月30日の直後から、人間が書いたような自然言語の生成能力に関心が集まり、米国統一司法試験の合格点を上回るだけでなく、人間の合格者の上位10%に入るスコアを獲得したとされています。また文章以外、例えば画像を生成するAIも急速に進化している状況です。大学教育においても、進化を続けるAIとの接し方は大きな議論をよんでいます。本稿では、AIの技術的な観点をいくつか選び、その観点に立って大学教育において我々、人間がAIとどう向き合うべきかを考えていきます。

さてChatGPTなど生成AIが文章を生成できる仕組みは、まずウェブから集めた大量文章から、ある単語の後にどんな単語が続くかを調べて、その単語同士の接続の仕方を確率的に表したものである学習モデルを作ります。そして利用者などから文章が与えられたとき、その学習モデルを通じて、その文章に続く単語を選ぶことを繰り返して、新たな文章を生成します。ここで注意したいことは、ChatGPTは単語と単語を繋いでいるだけで、生成している文章の意味を理解しているわけではないので、当然、間違っただけの内容を含む文章を生成することがあるということです。そもそもChatGPTなどの生成AIは知っていることと知らないことの区別がない、つまり生成AIには無知の知がないのです。かつてソクラテスは無知の知がある者は、ない者より賢いと指摘しましたが、その意味ではAIよりも人間の方が賢いことになります。その人間の有利性を維持するためにも、大学教育において学生に知識を教授することも大切ですが、同時に学生自身が自分に欠けている知識を認識することも大切になってきます。

その結果、ChatGPTなどの生成AIは間違っただけの内容を含む文章を生成しがちです。AIを使う場合、間違いが比較的少ない用途に使うか、人間がAIの間違いを発見・修正できる場合に利用すべきとなります。前者については要約や校正などの用途、文章の生成に必要な情報が与えられている場合は比較的错误が少ないです。後者については、自分が知らないことの内容を生成を求めると、利用者がAIの間違いを発見・修正ができなくなることから、ChatGPTなどのAIの利用では、利用する側である人間にはAIよりも高い知識が要求され

ることになります。大学教育では学生が課題レポートにおいてAIを利用することの是非が議論されていますが、学生がレポートの課題に関する知識があり、文章執筆を支援するためにChatGPTを使うケースと、知識がないのでChatGPTに頼るケースでは状況が違い、それぞれの是非以前に分けて考えておくべきです。

ChatGPTの登場により注目された言葉に、プロンプトエンジニアリングがあります。これは利用者によるChatGPTなどのAIへの入力、つまりプロンプトを工夫すると、AIの出力の品質が上がるという技巧です。ChatGPT内部では学習モデルがプロンプトに対応した文章を適切に生成するためにプロンプト改変・拡張する仕組みが多用されており、それを利用者が行っていることとなります。プロンプトエンジニアリングという技巧に本質的な意味があるかは疑問ですが、プロンプトの文章の書き方がAIの生成内容の品質に大きく影響するのは事実です。ChatGPTなどのAIの登場で、文章を書く能力、文章力は不要という考え方もありますが、むしろ逆で、適切なプロンプトを作り、AIを活かすには高い文章力が必要となるはずですが。

ChatGPTなどの生成AIは機械学習、その中でも深層学習と呼ばれる技術を利用しています。機械学習によるAIは、その学習モデルにおいて学習用データからルールを獲得して、そのルールを適用することで未知のデータに対応できるという、汎化と呼ばれる特質があります。例えば日本語と英語などの異なる自然言語の文章を訓練データとして学習モデルを構築すると、各言語の文章を生成できるだけでなく、言語間の関係、例えば相違な言語において共通する性質を見だし、それを各言語の文章生成に利用できます。ですから、日本語の文章生成は日本語の訓練データだけに頼っているのではなく、英語などの他言語向けの学習も活かしています。大学教育において、履修科目の選択は学生に任されていることも多く、各科目は独立して行われることが大半ですが、学生は相違な科目間に共通の知見などを見いだす能力が求められるはずですが。

ChatGPTなどの生成AIは所詮、道具です。AIという道具の特性と限界を理解した上で、道具を上手にコントロールして使うことが大切です。