「2.1.5 データサイエンスに基づく問題解決」に関するコラム
― 生成 AI ―

テキスト「2.1.5 データサイエンスに基づく問題解決」のコラム「人工知能（AI）」（p.107-）では，“人のような知能を持った機械の実現”に向けてどのような試みがなされてきたかについて，大きな進展の見られた３回のブームとして，主に歴史的な視点で紹介している．第一次と第二次の技術は一般に広く使われるほどの力を持ち得なかったものの，第三次のブームでは深層学習に基づく生成 AI（Generative Artificial Intelligence）を生み出し，社会のあらゆる活動に大きな影響を与えている．テキストに盛り込めなかった生成 AI の現時点（2024 年 2 月）での概要と関連用語を，このコラムで述べる．

○概要と発端

生成 AI が大きな影響力を持つのは，その汎用性の高さと使いやすさにある．生成 AI に対して質問や指示（プロンプトと呼ぶ）を出せば，豊富な知識を持つ多芸な人間のように，あらゆる話題に対して，文字だけでなく画像・音声・プログラムコードなど多様な方法（マルチモーダルと呼ぶ）で，適切な回答を返してくれる．

最初に注目されたのは，2018 年 6 月に発表された米 Open AI 社の自然言語モデル GPT（Generative Pre-trained Transformer）に基づく対話型 AI サービス ChatGPT（2022 年 11 月公開）で使われてからである．どのような話題であっても，適切な指示（文字だけでなく画像等を含む）で依頼すれば，自動で応答を生成して会話形式で利用できる．同様のチャットシステムとして，Microsoft 社の Bing（Bing Chat は Microsoft Copilot に名称変更）や Google 社の Bard（その後 Gemini に名称変更）が続いている．

さらに，チャットだけでなく API（Application Programming Interface）経由で，各種 Web アプリケーションからも利用できるようになった．その強力な機能のため，一時のブームで終わることなく，なくてはならない実用機能へと，応用範囲が急速に広がっている．このため，パソコン OS として最も広く普及している Microsoft 社の Windows において，標準キーボードの「アプリケーションキー」が「Copilot キー」へと約 30 年ぶりに置き換えられ，ボタン一つで生成 AI による支援機能を起動できるようになった．

○構成と特徴

このような機能が実現できたのは，深層学習を基本としながらも一段と発展させ，２段階の学習を取り入れたことによる．従来の AI では用途ごとの知識を学習させていたのに対して，まず大量のデータを使って汎用性の高いモデル（大規模言語モデル（LLM: Large Language Model）と呼ぶ）を作り，その後用途ごとの個別学習をさせる．これにより，一度 LLM を作れば，その後の学習の手間を少なくでき，短時間・低費用で多機能な高度システムを実現できる．さらに，画像処理・言語変換などの関連機能と連携させることも容易なので，高度な拡張性を持たせることもでき，文書だけでなく画像や音声・動画・ソースコードなどを生成できるマルチモーダル化も可能になった．

生成 AI の出現は，人と AI の協調作業において対話を通した思考や行動の高度化が可能になり，問題解決能力を大きく高めてくれる．ただ，LLM の構築やマルチモーダル化

には，大量データや高性能システムを準備しなければならず，だれでも作れるわけではない．とくに，並列処理に優れた GPU（Graphics Processing Unit）の大量利用が必須で，資金や人的資源の豊富なクラウド環境での利用が欠かせない．

さらに，使っている生成 AI がどのような知識を使って学習しているかを理解しておかないと，回答にもっともらしい嘘（ハルシネーションという）が含まれていたり，差別的・暴力的な偏見に満ちた意見だったり，著作権侵害や個人情報保護違反などの問題が生じる．使っている生成 AI について，不適切な回答をどのように抑制しているのか，個人情報を学習せず削除できる機能を持つのか，誤情報の出力を削減するようできているのかなど，仕組みが説明され公開されていることが必要である．これらの仕組みと考え方を十分に理解し，回答が適切かどうか使う側で判断し，責任を持てるようにしなければならない．

〇活用の広がり

生成 AI は，人間が必要とする問題解決能力のすべての場面で活用可能な，高い汎用性を持つ．文書・画像・音声・動画などさまざまな種類の情報を生成できことから，多様なコミュニケーション場面で利用可能である．

まず文書処理を取り上げる．適切なプロンプトを準備することで，報告書の作成，アイデアの具体化，問い合わせへの自動応答，文書の要約整理，楽曲の制作など，これまで人間の手で行っていた文書関係作業を大幅に効率化できる．生成できる文書には，プログラミング言語や Web ページのソースコードも含まれ，業務の課題解決に利用できる．さらに，回答に自分では思いつかなかったアイデアを含んでることもあり，思考の範囲を広げ深めることもできる．開発や設計に利用することで，高度な創造性を発揮する作業にも活用されている．

文書に限らず，画像の生成でも大きな進展が見られる．画像に関する技術として，画像を生成する拡散モデル（Diffusion model）と，画像と文章の組を学習したマルチモーダル基盤モデル（CLIP: Contrastive Language-Image Pre-training）とが重要である．拡散モデルは，画像にノイズを加えてランダム化しノイズだけにする拡散過程と，ノイズを取り除いて元の画像に戻す逆拡散過程からなる．これらの過程について，単語や文章の意味・関連性・文脈などを表すテキストと画像との関連付を，CLIP として学習する．プロンプトに含まれる指示を元に CLIP で制御すれば，指示に沿った多様な画像を生成できる．訓練データがなくても，少ない計算資源で高精度な任意の画像を生成できるため，広く使われている．

動画生成についても進化している．上で述べた画像生成により，高精度のフレームを複数枚作り，これらを補完しつつ連結すれば動画になる．ただこの方法では，フレーム間のつながりに不自然さが生じやすく違和感が残ることがある．これに対して，2024 年 1 月 24 日に Google 社の研究チームの発表した動画生成 AI "Lumiere"は，動画全体のフレームを一気に生成するという手法により，フレーム同士が自然につながった違和感の少ない動画を生成可能である．さらに，テキストと画像から動画を生成できるほか，写真の一部を動画化したり，動画の一部分を指定して加工できるなどと，動画生成の自動化と高度化を実現している．

同社はこれを発展させて，2024 年 2 月 15 日にはテキストプロンプトから，現実的で想像力豊かな最長 1 分の動画を生成する AI モデル"Sora"を発表している．ただ，あまりに

も高精度なため，重要な安全措置を講じないとディープフェイク作成など悪用の懸念がある．そのため，レッドチーム（後述「〇弊害と対応」のレッドチーミング参照)）と協力して安全措置を模索しており，確認できれば一般公開されるという．

日常の学習においても，身近に有能な家庭教師がいるように，積極的に利用すべきである．学校で学ぶときには，集団のなかでのコミュニケーションを通した協働的な学びと，自身の能力を高める個別最適な学びとを，一体的に進めなければならない．この後者の学びで，仮説検証やアイデア出しの壁打ち相手に使うなど，生成 AI を使った学習は有益である．自分が今どの段階にいるかを見極め，何が不足しているかを整理し学習目標を定め，これにふさわしい適切なプロンプトを作る一連の操作が，良い学びになる．さらに，協働的な学びにおても，プロンプトの違いによる回答の違いを比較するなどすれば，生成 AI の活用を通して学びの高度化が可能となる．これは，望ましい出力を得るために，効果的かつ魅力的なプロンプトを設計できる新たな職業であるプロンプトエンジニアに通じる．

〇 AI 規制と国際ルール

これまで人間の手によって開発・制作されてきたコンテンツが，今後は生成 AI によって生み出せるようになった．正しく使えば有益な生成 AI も，人をだますためや，一定の方向に誘導したり，組織や個人を誹謗中傷するために使うこともでき，社会生活全般に大きな悪影響を与える懸念が生じている．このため，全世界的な AI 規制のルール作りが必要と考えられている．

国際的な議論の場として，主要七カ国首脳会議(G7)や国連主催のフォーラム"Internet Governance Forum(IGF)"などがある．G7 デジタル・技術相会合では，開発者から利用者まですべての AI 関係者が守るべき責務の概要を示した指針と，開発者向けに責務をより具体化した規範から成る国際ルールを示している．指針には「市場投入前に適切な措置を講じる」ことなど AI 開発者に求める項目が含まれるとともに，「AI 固有のリスクに関するデジタルリテラシーの向上」ことなど利用者に求める規範までが盛り込まれている．

AI 規制については，欧州連合(EU)が最も進んでいる．AI を規制する法（AI 法）の整備が進んでおり，リスクを４段階に分類し，危険度に応じた対応を求めている．個人情報保護に関わる GDPR（テキスト p.262 参照）のような法規制となるので，国際間のデータ利用に多大な影響が出ると予想される．

こうした公的な規制だけでなく，LLM を開発する企業が国などと連携して，安全性を担保するための規制ルールを作る動きがある．アメリカや日本では，一気に法律を作るのではなく，技術革新と安全性のバランスを重視した開発・利用にあっての，法的拘束力を持たない原則の公開から取り組んでいる．アメリカでは，ホワイトハウスの科学技術政策局が，大手開発企業の説明責任を追及でき，国民の市民権を保護するために，「AI 権利章典のための青写真」として，AI を用いた自動化システムを設計・使用・配備するときに考慮すべき５つの原則を示している．日本でも，人間中心で安全性と透明性を確保するよう，安全性やプライバシー保護「AI ガイドライン案」を公開し議論が進んでいる．

さらに個別の組織が利用ルールを制定していることもあるので，生成 AI 利用にあたっては，まず自分の属している環境を広く見渡してルールを理解しなければならない．その上で，ネット空間で問題となっているフェイクやヘイトスピーチにならないよう，これま

で以上に広範な注意が要求される．

〇生成 AI の機能拡張

生成 AI の作る回答に関して，その質を向上させるために種々の工夫がなされている．このうち，データの追加拡張と，オープンソース化および日本語処理を取り上げる．

生成 AI の基盤となる LLM を構築すると，その時点での利用した範囲のデータに基づいたシステムとなる．このため，学習時に存在しなかった最新情報を追加したり，組織固有のデータを学習させる仕組み（ファインチューニングという）が必要となり，このために探索拡張生成（RAG: Retrieval Augmented Generation）が使われる．プロンプトに含まれる情報に基づき，検索エンジンや類似性を判断できるベクトルデータベースを使って，関連する知識を参照可能にする（これを接地（ground）させるという）ことで，生成する回答の質を向上させる仕組みである．

平等で公正なシステムの作成には，民主的なプロセスが必要で，生成 AI の開発でも同様である．このため，オープンソースソフトウェア（テキスト p.82 参照）の考え方が参考になる．多くの人の目に触れることで，完成度を高め問題点の迅速な発見と修正できる性質は，価値判断の絡む生成 AI にも通じる．たとえば，米 Meta 社の LLaMa2（Large Language Model Meta AI 2）は，先発のクローズドソースソフトウェアに対抗するオープンソースソフトウェア LLM の代表格である．

多くの生成 AI は英語のデータで学習した LLM を使っており，日本語を使った場合に，質問者の意図を的確に伝えるため質問を工夫したり，得られた回答の理解での精度が悪くなるなど，問題が生じることがある．このため，日本のスタートアップや大手通信企業により，日本語による LLM 作成が試みられている．日本語を使うことで法律面や倫理面で日本にふさわしい学習データを準備でき，質問や回答の精度が高まり，回答も日本語で扱いやすくなる．さらに，必要なパラメータ量と計算資源を削減でき，安全で使いやすいシステムを安価に構築できる．

〇弊害と対応

生成 AI の構築と利用では，インターネット上の大量のデータを使った LLM 構築段階（パブリック）と，利用者の属する組織の保有するデータを組み入れる段階（プライベート），ならびに，得られた回答が利用者の考え方にとって問題ないかどうか確認する段階（パーソナル）という三つの段階を意識し，各段階にふさわしい対応する必要がある．パブリックでは，その LLM が，だれにより，どのように作られたかが公開されているかを確認し，プライベートでは，属する組織のデータやルールを確認し，それに従っているかを点検する．パーソナルでは，無難で平均的な説明ではなく，個性があり独創的なものになっているかが重要であり，回答の単純なコピー&ペーストは避けるべきである．

ただ，いかに注意深く使って悪意がなくても，意図しない悪い結果を生み出してしまうことがある．差別的な表現や政治的な偽情報を回答したり，偽画像を作ってしまう危険性（これを，埋め込まれた弊害（embedded harm）という）をはらんでいる．これを防ぐには，作る側の高い倫理観とともに，システムを広く公開し，多くの目で確認することが必要である．たとえば，レッドチーミング（Red Teaming）という手法（倫理的ハッキングとも呼

ばれる）では，専門家による攻撃・点検チームを作り，誤った情報や偏見・悪意を含むコンテンツなどの問題を研究する．このチームが，許可を得て各種の確認作業を行い問題点を洗い出し，システム改修に活かす積極的な手法である．この手法を広く公開した組織で行ったり，組織内にチームを作って行うベンダーもある．

また，生成 AI の多機能性は社会のあり方や活動に影響を与えており，その一つに仕事を奪うという批判や反発が起きている．速度や正確性において人間によりも優れているコンピュータの出現時も，生成 AI と同様に人の仕事を奪うという恐れから批判的な動きがあった．ただ恐れたり否定的にとらえるのではなく，今の時代に生きる我々は，生成 AI の本質を深く理解し，適切に使えるよう努力し，自分の仕事にどう活かすかという肯定的な対応が望まれる．

職業選択においても，自分の将来をしっかり考えて，何を仕事にするか選ぶ時代が来ている．自分の職業選択は，生成 AI が回答できる内容ではなく，各自が主体的に答えるべき本質的な問題である．職業に与える影響に関しては，多面的な視点からの多様な意見が公表されているので，各自で調べて見ると良い．このとき，フューチャーデザインという取り組みが参考になる．これはさまざまな課題に対し，その課題の影響がおよぶ将来世代の立場も踏まえて議論しようというもので，時間軸で視点を変えて自分の将来を考えるのに活用できる．

○生成 AI との深いつきあい方

生成 AI の出現により，人類の技術に関する向き合い方が議論になっている．典型的なのは，効果的利他主義と効果的加速主義との対立である．前者は，貧困問題の解決や将来世代への責任を果たすことを優先し，公平な態度や他者との協力などを第一とする立場である．一方後者は，資本主義システムを加速することで社会変革を起こすことを優先し，情報技術で加速させた自由市場こそが効率よく社会問題を解決するという立場である．功利主義的な思想と自由至上主義的な思想とのどちらが良いか，一方的に決定できるものではないので，直面している課題に応じて各自が深く考えて決定すべきである．

これを考えていると，この対立の奥には，世界は一神教の元で因果関係に基づき直線的に変わっていくという西洋的世界観と，多神教の元で循環的・確率的に変化する東洋的世界観とがあることに気づく．どちらの世界観が正しいかという絶対的判断ではなく，どんな世界を実現したいかという価値観と，そのとき守るべき倫理観とを明らかにして，説明責任が果たせるように，深くつきあうことが一人一人に求められている．

これらの主義の対立や世界観の違いは，人間同士のつきあいでも起きることである．生成 AI の利用においても，ぶれない自分なりの価値観と倫理観を持ち，質問力や理解力（総合的な学び取る力）が必要となることは変わっていない．コラム「人工知能（AI）」（p.107-）の最後で述べた「人ならではの検討を加え，必要なら異議を唱えたり，技術者ならシステムを修正する態度」は，生成 AI の出現があっても変わっていっておらず，むしろより深めることが必要と言える．