

《論文》

進化する生成AIを巡る諸動静について

綱川 菊美

進化する生成AIを巡る諸動静について

綱川 菊美

和文抄録：本稿は、大量のデータを学習し、各種そのパターンを理解することで、文章や画像、音声等、様々なコンテンツを生成するAI（Generative AI）について、その驚異的な進展状況、世界に与えた衝撃、及び必至な変革への対処要請等、トレンドを概括して記した。現在、世界の才能、巨額の資金が、生成AI関連R&Dに集結し、その成果を頻々とローンチし続けていることから、それらの社会への実装、また潜在能力の押し広がりを見越した規制やフェアネスモニタリングの重要性について指摘した。

キーワード：生成AI、ChatGPT、OpenAI、AI規制





はじめに

生成AIの進化、広がりが凄まじい。2022年11月30日にOpenAIが世に出した「Chat GPT」は、瞬く間に世界中のICT精通者、特にAIエキスパートの間に広まり、1年8カ月超を経た今日（2024年7月末）では、かなり一般にも浸透し身近なツールとなりつつある。当初、世界中の多くが、同種の生成AIソフト公表の度、直後すぐさまログイン・アクセスをトライするも、「Oops!!…」から続いて、e-mailアドレスまで「リフレッシュしてトライせよ」との対応パターンが、何度も数多く繰り返された模様である。「Chat GPT」は公開5日後、世界におけるユーザー数が100万人に達し、2か月後には1億人へと激増している。2024年7月の最新データによると、そのユーザー数は数億人、2024年5月時点における月間訪問者数は23億人超と驚異的な数値を示している。今日尚、日々新たな生成AIサービスはローンチされ続けており、世界を一新させるような優れモノも多いが、「Chat GPT」は今後もその成長は続くと言われている。

さて、生成AIソフトは、いずれも高い能力を有し（表1参照）、人々の創造性を刺激して、様々な高次利用を手掛ける達人を多く輩出している。そうした中、驚嘆、束の間の呆然を余儀なくさせられた向きも少なくなかったとの報告が多い。プロの漫画家が、4コマ漫画作成を最新の生成AIソフトで試みた動画をみると、指示命令、提案等プロンプトに対する成果への反応は、大同小異、驚嘆の後、「仕事がなくなる！」と自嘲気味でエンドとなっている。それほどに、人間としての最後のよりどころ、創造性への生成AIによる侵食は、致命的に思えたようである。

アメリカはBIG TECHによる最新版生成AIソフト群は、それぞれ固有の背景特性を基にアピールに余念がないが、先行したOpenAIの新バージョン「Chat GPT」も、その都度前段のもの課題をかなりクリアし、そ

表 1. 生成AIで実現できることの例

		知的業務への活用		創造業務への活用	
	テキスト	添削、校正	要約、解釈	対話、プレスト	論点の洗い出し アイデアの提案
	プログラム	添削、校正	解釈	プログラムの生成	新規アプリの 開発・生成
	画像／動画	画像の不純物 除去	高解像度化	人物、背景 スタイル変換	3D画像生成 画像間の映像生成
	音声／音楽	ノイズ フィルタリング	音質高品質化	声色や抑揚の変換	BGMの生成、作曲

出所：PwC Japanグループ「2023年AI予測調査 日本版」PwC Japanグループ、2023年6月。

<https://www.pwc.com/jp/ja/knowledge/thoughtleadership/2023-ai-predictions.html>

の仕事ぶりは優秀なスタッフのものと同様以上へと飛翔している。故に、様々な分野の達人からも高い評価を得て、急速にその利用が広がっている。だが同時に、信頼性、安全性等の点から、課題の噴出も、懸念と抱き合わせで積み増されるに至っている。しかし、2024年も半ば過ぎの現在では、生成AIの進展に伴い、後追いあるいは並走する形で、これらの課題に向けたチャレンジングな技術、構想が奏効し始めている。

達人たちの生成AIソフトのユースケースに目を凝らすと、複数のデバイスを手にしてマジックの如くにそれらを操り、数人分、数時間分の仕事を、当初来、かなり高効率に処理している。多くの新しいバズワードが頻出する説明、ガイダンスには、戸惑いを覚える向きが一般にはまだ多いが、その仕事ぶり、成果の出来映えが相当なレベルであることの確信から、各層において、多くの興味、関心を刺激し、反響が、新たなユーザーを生んでいる。

デジタル化によりテキスト（文字）、画像、音声等の変換が自在となったことを利した、上級ユーザーたちのソフトや入出力法の多様な使い分けは、驚異的に超高効率な仕事ぶりを実現し、各分野第一線のプロフェッショナル達にすら、複雑な戸惑いや不安を覚えさせ、今後に向けた対応を急迫するようになっていくことが特記される。

とにかく、生成AIソフトを用いた、種々のストーリー展開性のある画像の作成プロセスを、動画配信ものでチェックすると、プロンプトに即応して、背景、表情が瞬時に次々と作成表出され、短時間で完成に至ることが明証されている。これら興味深い展開の中には、継続性の遮断というエラーもあったが、多くの人的作業の代替を見事にこなし、その仕事ぶりは圧巻であった。その他、種々の文書、また様々なコンテンツ加工関連作業への適用も、驚嘆、衝撃がここかしこに認められ、人々の関心、創造性を大いに刺激して、新たな画期的製品やサービスの発案、創出を促している例も出ている。

一方で、各国ないしEU等各圏域内、また国際レベルで、生成AIの急速な普及に向けた対策の検討が進められている。日本がイニシアチブをとった2023年5月の「広島コミュニケ」をベースに、そのビジョンに即した具体的な対処、方策の議論も諸々進展を見せ始めているが、他にも各界固有の対処要請、提訴に即応した緩急施策の明示が認められる。

以下本稿では、生成AIの威力、極めて高い応用可能性、またその進展を踏まえ、それらがもたらす広範な社会的影響を概観し、生成AIに向けた戦略の探訪、展望を図る。

1. 生成AIの急激な進展

生成AIの最前線は、俊英たちの凄まじい闘争の場となっている。新興勢の猛追、爆速成長もあり、主戦場の熱狂は正にAI大競争の時代の如き様相を呈している。ディープラーニング登場時をはるかに凌ぐインパクトで、広く社会に衝撃をもたらした「Chat GPT」だが、その後、世界のBIG TECHも新潮流を見極め、相次いでそれぞれの特性を踏まえた生成AIソフトを手掛け、公表している。「Chat GPT」の生みの親であるベンチャーOpenAIも、2022年11月末に全世界に向けて号砲を放った後、さらに進化を遂げたものをシリーズの如く世に放ち続けている。その急速な進展は、新たなAIの転機を、ICT達人ばかりか一般のユーザーにも強く印象付けている。当初来、「Chat GPT」の高度な対話能力の具備については大衝撃であり続けているが、識者のコメントは、その恩恵や脅威、課題への言及ついでに、多くが試用することで洗練度が増し、新たな適用等、創発可能性への展望をポジティブに語っている。そこには、強化学習が言語モデルに適用され、ユーザーからのフィードバックが、人間に固有の感覚や嗜好に、より一層即すよう、調整可能なアプリケーションとしてあることの評価もあろう。実際、多くが生成AIを利活用する効果は、その潜在力が大きく押し広げられていることに見取れ、畢竟今後の成長の継続にもつながっていくと推計される。

ここで、「Chat GPT」が如何ようにして進化を急速に遂げ、人々にとって有用なツールとして普及し始め、インフラ化が進みつつあるかをラフレビューしておく。

これまで検索作業については、多くが巨人グーグルの独壇場に依拠していた。「ググル」なる語が、若年層間では「検索」に代わる語として、目下日常用語化し定着しているほどであるが、これにはいくつものステップを踏み、さまざまな知見を駆動させての作業が必至である。しかし「Chat GPT」は、言葉という誰でもが日常使用している自然言語インターフェイスでのオペレーションが可能であり、これまで多くが苦手あるいは忌避していた種々の仕事を、巧みな問いかけにより、効率よく目指す成果につなげることができるようになっていく。但し「Chat GPT」シリーズの場合、学習データが過去のものであり、更新されたとはいえ期限があり、未来データは保持していない。そこに、AIが事実に基づかない情報を生成する現象、ハルシネーション(Hallucination)問題^(注1)が起きるリスクがあり、世間を騒がせた事例も少なくはないことから、種々対応策が検討されている。今日では、多様な生成AIソフトが日々ローンチされる中、リアルタイム情報を絡め込んだ回答を返すものも出ている。有料版Perplexity(対話型AI検索エンジン)は、先般の都知事選(2024年7月7日投票)において、リアルタイム情報を、大方既存メディアから掬い上げた上での回答返しをしていた。しかし、これら既存メディア等からの最新情報学習に規制がかかることになれば、その仕事は成立の可否を問われ、価値の減衰ないし消滅に帰すことになるだろう。

さて今日、我々がその圧倒的な利便性、いわば多言語対応でさらに広範囲かつ高度な仕事能力の発揮に衝撃を受け、魅了された「Chat GPT」は、同じOpenAIが2020年に公表した「GPT-3」の進化形としてある。その起源はカナダのヒントン教授率いるチームにより、2006年に登場した第3次AIブームのトリガー、ディープラーニングに遡る。生成AIソフト、中でも「Chat GPT」の急激な進展の軌跡については、2009年のグーグルによる「データこそ正義」というタイトルの論文発表、その後のビッグデータ旋風、翌2010年前半のAI向けGPU(Graphics Processing Unit:画像処理装置)の本格的開発、処理能力の飛躍的進展、同年後半におけるディープラーニング向け種々の半導体の出現、これら一連の成果を受けてハードウェアの伴走力が高度化したこと等が主要なものとして指摘される。さらには、生成AIの肝ともいえるべきグーグルによるトランスフォーマーの発明が2017年にあり、大量のデータを用いると精度が向上することが周知となり、2018年以降は、このトランスフォーマーを用いた高精度AIが相次いで誕生することになった。今日、誰でもがクリエイターになれるような多彩で優れた生成AIが、身近なツールとなり得たブレークスルーとして、これは銘記すべき卓抜した重要技術である。

AIに必要な要素としては、回答を導くためのモデル、そのモデルに学習させて最適化するためのデータ、そ

して計算を実行するためのハードウェアが指摘される。OpenAIは、2022年1月に公にした「Chat GPT」の兄弟、言語モデル「InstructGPT」に関する論文「Training language models to follow instructions with human feedback」^(注2)において、スケーリング則以外にもAIの性能を向上させる重要な要素があると説き、事前学習済みの「GPT-3」に対して、教師あり学習によるファインチューニング、及び人間のフィードバックに基づく強化学習の付加が、「InstructGPT」の開発を成功に導いたと明言している。2023年3月14日、Open AIは「GPT-3」、「GPT-3.5」の上位版として「GPT-4」を正式発表した。これは前段のものに比し、高度な出力精度、誤字・脱字の大幅減少、複数の指示を含む応用的プロンプトへの対応に加え、マルチモーダル化を殊にアピールした。また2023年11月6日には、最新版大規模言語モデル「GPT-4Turbo」を公表し、一度に対処可能な文章量（上限で約300ページ分の本相当）と知識の増強により、誰でもが自然言語で用途を特化した「Chat GPT」の作成が可能な新機能「GPT s」を披露した。2023年4月までの情報を取り込んで学習データの更新を図り、それら新情報を踏まえた回答をも可能にしている。他の改善点を見ると、特定のフォーマットを生成する等、プロンプトにより慎重に対応する必要のある高難度細密なタスクに対しても、「GPT-4Turbo」はこれまで以上に優れたパフォーマンスを発揮し、より良い聞き手としてユーザーに有益な成果の提供を可能ならしめている。また開発者がAPI（Application Programming Interface）を使用する際の料金が、インプットで3分の1、アウトプットで2分の1となり、開発コストの低減に資するようになっており、加えて、これまで面倒で悪評だったドロップダウンメニューに変え、「GPT-4Turbo」では、プロンプトでの画像要求に対し、自動的に画像処理ソフト「Dall-E3」を使用して回答対処するといった進歩が期待できるようになっている。適切なツール選択の自動化の実現である。

しかし、近時、出版社や著者等から、同意なしでの著作物利用についての異議申し立て論調が高まりつつあり、情報取得問題の解決が急がれている。概ね、運行監視と脆弱性の診断を踏まえ、ネガティブを補完する手法の模索は、不断に行われているとされているが、難題は尽きそうにない。

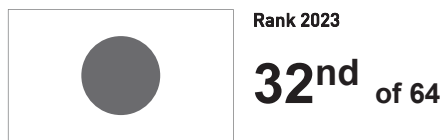
OpenAIは、「全人類が汎用AIの恩恵に与えられるようにする」ことをミッションとして、2015年12月にスタートしているが、設立時来、世界トップの頭脳集団として注目され、圧倒的な成果を誇って、AI界を牽引してきた。その威力は、「フォーチュン500」企業の92%による製品の利用、約200万人の開発者がOpenAIのAPIを使用していること、またジャイアントテックを刺激して、彼らの生成AI総参戦を促し、ついには、誰でもがオンデマンドでスーパーパワーを手中にすることを目指しているようである。だが、今後も一強としての威光をキープし続けられるか、新興勢の爆速成長との競合の行方が注視される。

2. 生成AIの利活用と課題

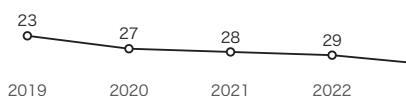
組織として、技術トレンドをキャッチアップしていく眼力は重要な資質だが、それには勘所に優れた人材が欠かせない。種々の必須リソースを探って一步を踏み出し、先行を確実にするには、潤沢な資金、関連する有用情報に加え、トップの胆力も必須であろう。生成AIの如き常識を覆すような革命的超高性能ツールに対し、未だ黎明期とはいえ、取り組みの先延ばしや無頓着は問題である。だが、種々の調査は、日本の対生成AI取り組みが、総じて低調気味であると示唆している。IMD（International Institute for Management Development）による2023年度世界デジタル競争力ランキング^(注3)が示した、日本の年次ランク低位化（図1、表2-1、表2-2参照）を裏付け、実感させるものが少なからずアウトプットされている。未だ生成AIの利便性の学び、認識の不十分さから、活用可能領域への知見不足等で、導入、利活用辺りの段階での問題が浮き彫りにされている。確かにAIについては、ビジネス界や教育分野での導入が近年急速に進んでいるものの、業務の改革等への利活用の方は一部を除き遅々としているのが実態であり、具体的な成果の実感が必ずしも芳しくなさそうである。

東京都では、学校職員を含む約5万人の都職員が、2023年8月23日から「Chat GPT」を業務利用することを可能にし、仕事の効率化やサービスの向上を図った。にもかかわらず、同年12月19日の利用実態調査結果報告によると、約10%、5,000人の利用に止まっていたことから期待外れであったと評している。多くの同種調査

図1. **IMD**
Digital Competitiveness Ranking
JAPAN



Overall Performance



Peer Group Ranking - Asia - Pacific



出所：IMD Digital Competitiveness Ranking 2023

https://www.imd.org/news/world_digital_competitiveness_ranking_202311

表2-1. IMD世界デジタル競争力ランキング2023年

順位	国・地域名
1 (2)	米 国
2 (6)	オランダ
3 (4)	シンガポール
4 (1)	デンマーク
5 (5)	ス イ ス
6 (8)	韓 国
7 (3)	スウェーデン
8 (7)	フィンランド
9 (11)	台 湾
10 (9)	香 港
	⋮
19 (17)	中 国
20 (16)	英 国
	⋮
23 (19)	ド イ ツ
	⋮
27 (22)	フ ラ ンス
	⋮
32 (29)	日 本

注：調査対象64カ国・地域

()内は、2022年順位

出所：同 図1

表2-2. IMD世界デジタル競争力ランキング2023年
指標からみる日本の強みと弱み

[64カ国中上位10カ国に入った指標]

因子	サブ因子	指標	順位
知識	人材	教育評価（PISA-数学/15歳対象）	5
	訓練・教育	高等教育での教員一人当たりの学生数	3
	訓練・教育	高等教育修了率（25-34歳人口比）	6
	訓練・教育	高等教育の学位を持つ女性	6
	科学的集積	研究開発への公的支出（対GDP比）	6
	科学的集積	全特許に占めるハイテク関連特許の割合	6
	科学的集積	教育・研究開発用ロボットの数	6
技術	資本	IT&メディアの株式時価	10
	技術的枠組み	無線ブロードバンド普及率	2
	順応性	市民の行政への電子参加	1
将来の準備	ビジネスの俊敏性	世界での産業ロボット供給	2
	IT統合	ソフトウェア著作権侵害	2

[64カ国中下位10カ国に入った指標]

因子	サブ因子	指標	順位
知識	人材	上級管理職の国際経験*	64
	人材	デジタル/技術的スキルの可用性*	63
技術	科学的集積	女性の研究員	57
	規制の枠組み	企業での外国人雇用と移民法制*	62
	順応性	スマートフォン所持率	55
将来の準備	ビジネスの俊敏性	機会と脅威に対する企業の対応*	62
	ビジネスの俊敏性	企業の俊敏性*	64
	ビジネスの俊敏性	ビッグデータとアナリティクスの活用*	64

注：*は、エグゼクティブ調査の指標

出所：同 図1

とほぼ似た結果であり、何にどう使ったらよいのかわからないとの感想、意見が多かった背景には、やはり「Chat GPT」への習熟度が低いことが指摘される。サポートやトレーニングの不足、それ故か、操作が複雑でわかりにくい、カスタマイズが困難、応答の質に一貫性がない等、利用に消極的な理由はどの分野でも共通している。利用頻度が高いユーザーは、ICT達人でもあり、「Chat GPT」の課題指摘が、他の低使用頻度層とは異なり、高次元のものが上位に付置されている。自然言語処理での検索、分類、抽出、そして最難度の生成までが実用レベルに到達して来ているといった認識が広まり、浸透して行かない限り、リテラシーの格差が様々な悪影響を社会に及ぼすことになると懸念される。漸次、生成AIの先行ユーザーによる活用の実例に学び、イメージ学習の機の増大を利し、格差リスクの低減を図って、革命的なツールの恩恵享受が必至の時代になりつつあることの認識が広まることが望まれる。

技術の進歩が速度を増してビジネス環境が激変する中、世界レベルで、既に生成AIありきの経営に向けた転化は始まっており、職業の淘汰、リスクリングの推奨が勢いを増している。これまで、技術の恩恵は利便性と豊かさの実現に象徴、集約されてきたが、AIの急速な高次元での進歩は、人間のあり方や社会システムの根本的な問い直しを急迫し、労働や学びの再定義へと我々を駆り立てている。人類は新たな価値観を創出しなければならない岐路に立たされているといえよう。

3. AI規制法に向けての動き

今日では、アメリカが開発した革新的な技術を、追従する国々が学んで模倣し、その功罪を実体験、俯瞰してヨーロッパが規制をかけるといったような構図が見取れる。特にICT、AI関連については、アメリカのBIG TECHの独壇場や生成AI先行群の威力の影響の大きさを憂慮し、EUは包括的な規制を長きに渡って激しくかつ慎重に熟議してきている。そして、現地時間2023年12月8日夜に、EUは世界に先駆けて、生成AIを含むAIを包括的に規制する「AI法」を大筋合意するに至り、2024年3月13日には、世界初の包括的なAIに関する規制となるEUのArtificial Intelligence Act (AI Act, AIA)を議会の採択を経て最終的な全容を明確にした。AI Actの最新版は、2024年4月16日にEU議会から公開されたが、様々なリスクを懸念し、有害な革新の防止を期してはいるものの、AI開発の上流工程における革新を阻害する可能性を残している。最終的な文言の合意を経て、当該法案がEU議会において採択を得たものの、自縄自縛で投資家離れを来すとの批判もある。革新的な企業の主導が、市場を拡大してきたシリコンバレー型アプローチへの依存を脱し、独自の未来を切り開く礎となる画期的な法とは現時点ではなり得ていない恨みがある。

しかし、EUが合意した「AI法」は、デジタル空間のみならず、リアル空間にある種々のAIを使用した製品やサービスにも影響することになる。国境のないグローバル・デジタル経済下、この法律は外国法でありながら、外国法ではないとの言に、慎重に向き合い熟考対処する必要がある。今般の先進的な「AI法」は、「容認不可リスク」については禁止とし、対象の広い「高リスク」に分類されるものに対しては、義務の順守方法の策定を標準化機構に委ねる等、リスクベースの手法をとっている。ハードローとソフトローとの組み合わせにより、難題、難航のクリアーを図っている。

社会的影響の大きい生成AI等「汎用的AI」に特化した規制については、提供者に透明性の確保やリスクの評価・軽減等の特別な義務を課して、自主的な取り組みを強制力のある法律に取り込んでいくことにしている。

当該法は、強い影響力を有する特定のプラットフォーム等の事業者に対し、情報の開示を求めているEUのDSA (Digital Services Act) に類似しており、従前来のEUデジタル政策の延長線上にあり、DSAと並走することになる。今後、技術標準や行動規範の策定が進められ、AI法の義務化の具体化による実効性が期待されている。

さて、AIを主導してきたアメリカにおける規制作りに目を転じると、多民族、移民社会ならではの、また覇権を誇る国家ならではの問題が、生成AIのリスク問題に重なって反映されている。民主主義を脅かす要因となるフェイクニュースやその拡散、雇用や格差等、社会の歪みを惹起することへの加担、助長等、国民に深刻な

リスクとなり得たり、国家の安全保障にも影響するような技術開発をする事業者に対し、2023年10月、包括的大統領令が発せられた。開発時点での国への通知義務、また安全性テスト結果の提示義務等、最新AI技術の信頼性、悪用の未然防止等に向けての規制が明確にされた。アメリカ議会では、2023年度、50本以上のAI関連法案の提出があったようだが、中国の台頭やその他今日の世界情勢を鑑み、過度な規制への警戒感が認められる。だがアメリカでは、グーグルを筆頭にAI開発事業者7社が諸般のリスクに対し、既に自主的に対策を講じており、法的な一定規律の設置は、それらの後押しとなろう。

日本においては、これまで総務省や経済産業省によるAI関連指針はあった。だが、抽象的な内容であったため、実情は参照効能を逸しており、海外からは規制の緩い国と評されていた。しかし、2023年12月21日の第7回AI戦略会議では、事業者向け指針の内容を充実させ、具体的な点検リストの素案を明示した。AIの基盤モデルを開発する「開発者」、AIを組み込んだサービスを提供する「提供者」、事業でAIを利用する「利用者」の3主体分類をベースに、各々が順守すべき細則、及び全事業者向けの10の指針を明らかにしている（表3参照）。日本国内はもちろんEUやアメリカ等のAI規制に関連する資料を博搜し、有用事柄の詳解を束ねたものをタタキ台にして提示された指針は、事業者の順守を要件とし、法的拘束力は持たない。リスクの未然防止基準を示すのみでは、AI諸般の問題の防止、特に安全性の確保には限界がある。

そこで、指針の実効性担保のため、制度の整備や、安全性に関する情報開示の仕組み作り、また偽情報対策として、AIによる作成コンテンツの明示義務化等の検討を図り、世界のAI法制化動向に即していくとしている。2024年に入って以来、生成AIを含むAIに関する種々の検討会の立ち上げや、既存委員会等からの中間報告が出されているが、未だ作業の過程にある。AI関連技術の急激な進歩を眺みながら、様々なステークホルダーへの配慮必至の規制の検討は百家争鳴、熱い議論の成果が待たれる。

サイバー空間における覇者の巨大権力に対し、近年ヨーロッパでは、デジタル立憲主義なる語が使用され始めており、デジタル空間の健全性確保にコミットを深化させている。権力制限の射程を広げて、立憲主義的な価値による制御を図ろうとしている。世界が早晚向かうことになろうAI規制の収斂先への効用は如何ある？と思いが巡る。

終わりに

今日の世界経済では、イノベーションを真っ先に実現した企業が、その後も市場で大きなシェアをキープする傾向が目立つ。しかし、イノベーションは、必ずしも全くの無からの創造とは限らず、既存技術の組み合わせ、融合から、可能な限りこれまでにない新規のものやサービス等を生み出すことでもある。セレンデピティ、プロダクトイノベーションにプロセスイノベーション、これらが、生成AIの進展には、その貢献度、影響度の大きさに違いこそあれ功あったと解される。また近時、性能そのものよりも、多くが使っているから便利といったネットワークの外部性が影響して、利用者が増え続けている財やサービスが多くなりつつあることが付記される。世界各国の企業は業種を問わず、生成AIの最適利用を図って、先行者としての利益確保を目指し、従前以上に革新的な技術、製品、サービスの開発に取り組む姿勢が強まっている。日本では、近視眼的なコストカット性向が他に比し強い傾向があるせいか、経済が縮小均衡化してきた経緯から、新時代を牽引するイノベティブな技術への投資が全般に低調ぎみであった。世界レベルでデジタル技術の進歩がスピードを増している今日、そのコアたる生成AIの急激な進展は、多分野に影響し、現に第4次AI革命と称される状況を生み出している。急務の課題は諸々あるが、詰まるところ、革新的なイノベーションこそが持続的な成長の源泉であるという問題意識を改めて共有し、それをキープし続けて、万般、生成AIを初め革新的な技術に対する直言の大胆な見直しが肝要になっているといえよう。

生成AIの多用が行き着く先は、「小さな箱庭文化」となるのではとの展望評が文化関係者から出ているが、一方で、インターネットの急速かつ多彩な発展史に鑑み、生成AIの潜在力の解き放ちが、新たな世界をたぐり寄せるとの言も多見される。ともあれ、知の探究のスピードと広がりをも助長している革新的技術、生成AIについ

表3. 共通の指針及び主体ごとに重要となる事項

	共通の指針	共通の指針に加えて主体毎に重要となる事項		
		AI開発者 AI Developer	AI提供者 AI Provider	AI利用者 AI Business User
1) 人間中心	① 人間の尊厳と個人の自律 ② AIによる意思決定・感情の操作等への留意 ③ 偽情報等への留意 ④ 多様性・包摂性の確保 ⑤ 利用者支援 ⑥ 持続可能性の確保	-	-	-
2) 安全性	① 人間の生命・心身・財産、及び環境への配慮 ② 適正利用 ③ 適正学習	i. 適切なデータの学習 ii. 人間の生命・心身・財産、及び環境に配慮した開発 iii. 適正利用に資する開発	i. 人間の生命・心身・財産、及び環境に配慮したリスク対策 ii. 適正利用に資する提供	i. 安全を考慮した適正利用
3) 公平性	① AIモデルの各構成技術に含まれるバイアスへの配慮 ② 人間の判断の介入	i. データに含まれるバイアスへの配慮 ii. AIモデルのアルゴリズム等に含まれるバイアスへの配慮	i. AIシステム・サービスの構成やデータに含まれるバイアスへの配慮	i. 入力データ、プロンプトに含まれるバイアスへの配慮
4) プライバシー保護	① AIシステム・サービス全般におけるプライバシーの保護	i. 適切なデータの学習 (D-2) i. 再掲	i. プライバシー保護のための仕組みや対策の導入 ii. プライバシー侵害への対策	i. 個人情報の不適切入力とプライバシー侵害への対策
5) セキュリティ確保	① AIシステム・サービスに影響するセキュリティ対策 ② 最新動向への留意	i. セキュリティ対策のための仕組みの導入 ii. 最新動向への留意	i. セキュリティ対策のための仕組みの導入 ii. 脆弱性への対応	i. セキュリティ対策の実施
6) 透明性	① 検証可能性の確保 ② 関連するステークホルダーへの情報提供 ③ 合理的かつ誠実な対応 ④ 関連するステークホルダーへの説明可能性・解釈可能性の向上	i. 検証可能性の確保 ii. 関連するステークホルダーへの情報提供	i. システムアーキテクチャ等の文書化 ii. 関連するステークホルダーへの情報提供	i. 関連するステークホルダーへの情報提供
7) アカウンタビリティ	① トレーサビリティの向上 ② 共通の指針の対応状況の説明 ③ 責任者の明示 ④ 関係者間の責任の分配 ⑤ ステークホルダーへの具体的な対応 ⑥ 文書化	i. AI提供者への共通の指針の対応状況の説明 ii. 開発関連情報の文書化	i. AI利用者への共通の指針の対応状況の説明 ii. サービス規約等の文書化	i. 関連するステークホルダーへの説明 ii. 提供された文書の活用と規約の遵守
8) 教育・リテラシー	① AIリテラシーの確保 ② 教育・リスキング ③ ステークホルダーへのフォローアップ	-	-	-
9) 公正競争確保	-	-	-	-
10) イノベーション	① オープンイノベーション等の推進 ② 相互接続性・相互運用性への留意 ③ 適切な情報提供	i. イノベーションの機会創造への貢献	-	-

出所：AI戦略会議第7回令和5年12月21日「資料1-3 AI事業者ガイドライン案」、総務省、経済産業省。

*一部、著者編集。https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/7kai.html、2023年12月24日check

ては、注視すべき重要事象であることは言うを待たない。

【注】

1. ハルシネーションとは、AIが事実に基づかない情報を生成する現象のことであり、会話型AIサービスの信頼性に関わる問題となっている。ハルシネーションは、学習に用いたデータとは異なる事実の出力をするケース「Intrinsic Hallucinations」と、学習に用いたデータには存在しないことを出力するケース「Extrinsic Hallucinations」とに2分類される。
<https://www.nri.com/jp/knowledge/glossary/1st/ha/hallucination>、2024年1月5日 check
2. 「Training language models to follow instructions with human feedback」については、以下参照のこと。arXiv:2203.02155 [cs.CLの] または arXiv:2203.02155v1 [cs.CL]
3. IMD世界デジタル競争力ランキングは、デジタル技術をビジネスや政府、社会における変革の重要な推進力として活用する能力と態勢を、国、地域ごとに測定し比較したものである。公的統計及びエグゼクティブの調査回答から収集した57指標の中から、54指標を「知識」、「技術」、「将来の準備」の3因子と9つのサブ因子にまとめて毎年分析している。日本におけるデータの収集は三菱総合研究所と経済同友会とが連携して行っている。以下参照のこと。
https://www.imd.org/news/world_digital_competitiveness_ranking_202311

【参考資料】

1. AI戦略会議第7回令和5年12月21日「資料1-1 広島AIプロセス資料について」、「資料1-2 AI事業者AIガイドライン案概要」、「資料1-3 AI事業者ガイドライン案」、「資料2 AI戦略会議の今後の課題、内閣府」。
https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/7kai.html、2023年12月24日 check
2. AI戦略会議第7回令和5年12月21日「議事要旨」内閣府。
https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/gijiyoushi7kai.pdf、2023年12月24日 check
3. AI戦略会議第6回令和5年11月7日「議事要旨」内閣府。
https://www8.cao.go.jp/cstp/ai/ai_senryaku/6kai/gijiyoushi6kai.pdf、2023年12月15日 check
4. AI戦略会議第5回令和5年9月8日「議事要旨」内閣府。
https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/gijiyoushi5kai.pdf、2023年12月15日 check
5. AI戦略会議第5回令和5年9月8日「資料3-3 ChatGPT等の生成AIの業務利用に関する申合せ（第2版）（案）」内閣府。
https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/moushiawase.pdf、2023年12月15日 check
6. IMD Digital Competitiveness Ranking 2023
https://www.imd.org/news/world_digital_competitiveness_ranking_202311、2023年12月20日 check
7. 『日経ビジネス2023 03.20 No.2183 ChatGPT破壊と創造』「特集ChatGPT破壊と創造：「PROLOGUE 百家争鳴、勃興する「創るAI」」14～17ページ、「PART1対話型AIの威力に個人も企業も夢中」18～23ページ、「PART2 ChatGPTの舞台裏に4つのブレイクスルー」24～27ページ、「INTERVIEWS 世界の頭脳が語る「革命の本質」」32～35ページ」、2023年3月20日発行、日経BP社。
8. 『日経ビジネス2023 11.27 No.2218 同僚は生成AI 破壊者か期待のエースか』「特集同僚は生成AI 破壊者か期待のエースか：「PART1 相次ぐ盗用に解雇 問われる人間の役割」10～15ページ、「PART2 先進企業は既に転換 生成AIありきの経営」16～21ページ、「COLUMN 米テック経営者直伝「私はAIとこう働く」」22～23ページ、「PART3 3大トレンドが告げる生成AI大競争時代」24～29ページ、「PART4 新興勢が爆速成長 生成2023年AIの旗手25選」30～33ページ、「編集長インタビュー 澤田純 NTT会長 AIによる分断、哲学が止める」34～37ページ」2023年11月27日発行、日経BP社。
9. PwC JAPANグループ「2023年A I 予測調査 日本版」PwC JAPANグループ、2023年6月。
<https://www.pwc.com/jp/ja/knowledge/thoughtleadership/2023-ai-predictions.html>、2023年12月20日 check
10. 『知識資本の時代 世紀末大転換と激動の21世紀を診る』「第7章 知識労働と知識労働者」149～174ページ、「第8章GAFAMの深層を探る—有力諸説の検討（その1）」175～196ページ、「第9章GAFAMの深層を探る—有力諸説の検討（その2）」197～220ページ、関下稔 著、2023年6月10日初版、晃洋書房。
11. 『日経ビジネス2024 05.27 No.2242』「特集 瀕死のインターネットAIがばらまく嘘と悪意」8～29ページ、2024年5月27日発行、日経BP社。

Various Impacts of Evolving Generative AI

Kikumi TSUNAKAWA

Summary

This paper summarizes trends in generative AI, which learns from large amounts of data and understands various patterns to generate many different contents such as text, images, and voice, and discusses its phenomenal progress, the impact it has had on the world, and the inevitable calls for change. Currently, talent and huge amounts of money are being concentrated in generative AI-related R&D, and the results are being launched frequently, so the paper points out the importance of regulation and fairness monitoring in anticipation of their implementation in society and the expansion of their potential.

Key Words: Generative AI, ChatGPT, OpenAI, AI Regulation