

Title	生成系AIシステムがイノベーションに与える影響に係る考察
Author(s)	市川, 類
Citation	年次学術大会講演要旨集, 38: 11-15
Issue Date	2023-10-28
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/19280
Rights	本著作物は研究・イノベーション学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Research Policy and Innovation Management.
Description	一般講演要旨

1 A 0 3

生成系 AI システムがイノベーションに与える影響に係る考察

○市川 類 (一橋大学)
ichikawa@iir.hit-u.ac.jp

1. 問題意識

近年、ChatGPT を代表とする対話型の生成 AI システムが、世界中において爆発的に普及が進みつつある。この対話型の生成 AI システムは、人間の知的活動の基盤である言語を操るものであり、特に人間からの質問・入力（プロンプト）に対し、人間が作成したような文章として、適切な回答を行うことが可能となっている。その際、人間が同システムとこのような対話をするることによって、イノベーションの創出に役に立つとの指摘が多く識者によってなされている。

しかしながら、このような対話型の生成 AI が、イノベーションのプロセスの中において、具体的にどのような役割を果たすのか、また、その限界はどこかなどについては、必ずしも十分に整理・考察されていない。また、人類社会のイノベーション活動に対する AI（機械）の大幅な関与は、逆に、「なぜ人間はイノベーションを起こすことができるのか」という、社会におけるイノベーションの本質に係る考察が改めて必要になる。

このような問題意識の下、本稿においては、イノベーション論とそれに関わる人類の発展の歴史の観点から、対話型の生成 AI が社会におけるイノベーションにおいて果たす役割及びその限界を改めて整理するとともに、イノベーションにおける人間の本質的な機能に係る論点を整理することを目的とする。具体的には、まずは、次章において、これまでのイノベーション論とそれに関わる人類の発展の歴史に関して整理を行う。次に、第三章において、生成 AI の技術的な仕組みと限界について整理した上で第四章において、生成 AI との対話によるイノベーションの可能性についてその役割を整理する。また、第五章では、生成 AI に係る「再帰性の呪い」の議論を踏まえ、生成 AI と比較して、人間がイノベーションに果たす本質的な機能について考察する。

2. イノベーション論の整理と人類におけるイノベーションの歴史

(1) イノベーション論におけるイノベーションのプロセス

イノベーションとは、シュンペーターによると、「新結合の遂行」による経済社会の変革である[1]。その際、シュンペーターは、当時、モノと力の「新結合」を念頭においていた¹が、現代では、モノ、力、データの結合体系（ルーティン）に係る技術・知識に関して、その融合・新結合を図ることによって、モノ、力、データに係る新結合をもたらすような新たな技術・知識の創出することと整理することができる。また、イノベーションにおける「遂行」とは、現代の経済社会システムにおいては、社会ニーズに対応することによって新たな付加価値を生み出し、経済社会の変革をもたらすような「新結合」（新ルーティン）を実現し、普及することが目的となる。

このような観点から、イノベーションのプロセスにおいては、理念的には、以下の二つの段階が存在する（図 1 参照）。ただし、いずれの段階も、試行錯誤を通じて最適解を求めていくプロセスであり、両段階は互いに関連し合うことになり、必ずしも明確に分けられるものではない。

- ・ 「新結合」（アイデア）の創出段階（A 段階）：
それまでに社会において蓄積されてきた既存の技術・知識基盤をベースにしつつ、その多様な組み合わせの中から、新たなアイデアの創出する段階。その際、特に、Serendipity（偶然の産物）が重要な役割を果たすこともある。このようなアイデアの創出にあたっては、B に記す目的に照らしつつ、組み合わせに係る試行錯誤が必要となる。



図 1：イノベーションのプロセス

¹ これは、当時においては、製造業におけるエネルギー利用がイノベーションの中心であったためであり、現代のようなイノベーションにおける知識の重要性の増大と、情報化・デジタル化の進展の影響が必ずしも考慮されていなかったためものと考えられる。

- ・ 新結合（イノベーション）の「遂行」段階（B 段階）：付加価値の創出などを目的として、社会ニーズにあった新結合を特定し、その最適解となるもの²を目指して試行錯誤を行い、遂行していくという段階。これらには、実証実験の実施やアルファ版の製品・サービスの発表などの試行錯誤のプロセスも含まれる。

これらの段階に関し、イノベーションを促進する観点から、A 段階においては、様々な知見を有する多く関係者とのコミュニケーション等を通じたアイデア創出が重要であり、このためクラスター、ネットワーク、エコシステムなどのイノベーションシステムの在り方について議論がなされている。また、B 段階においては、市場メカニズムを含め、社会・経済におけるインセンティブ制度や組織のマネジメントの在り方などについて議論がなされてきている。

なお、科学研究に係るプロセスについても、このイノベーションのプロセスと同様、原則として、試行錯誤を通じて目的に向けて新結合が進行するプロセスであるとともに、イノベーションに係る全般的プロセスの一過程として共進化しながら進展するプロセスであると位置づけられる。ただし、科学研究の場合は、一般的には、現実世界の理解を通じて、「真理」と呼ばれるような理論・法則の構築を探索することを目的とするため、以下の2点において、通常のプロセスとは異なる。

- ・ 既存の知識をベースにして、仮説（アイデア）を検討しつつも、原則として、社会（マクロ世界、社会・人間、ミクロ世界）などからのデータ取得（観察、実験）とその解析に基づくその検証という試行錯誤プロセスを伴うものであること（特に A 段階）。
- ・ 経済的な付加価値の創出というよりは、むしろ、人間の言語で簡易に表現されるような新たな理論・法則を構築し、それがピア（関係者）の中で評価されることを目的とすること（特に B 段階）。

（2）人類・社会の進化としてのイノベーションとデジタル・AI の役割

このような現代の経済・社会におけるイノベーションのプロセスに関しては、生物の進化プロセスとのアナロジーで考える見方もある。具体的には、経済進化学では、生物が、遺伝子交換によって新たな遺伝子変異が生じる中、環境に適応した生体（遺伝子）が適者生存するというプロセスを通じて進化すると同様、経済においては、技術・知識が交換される中で、新たなルーティンが形成され、そのうち、経済・社会環境に適合したルーティンがドミナントになっていくプロセスがイノベーションであると位置づける（例えば、[2]）。

また、人類史においては、生物から人類への進化の過程において、遺伝子変異による適者生存から、イノベーションによる適者生存のメカニズムに発展してきたことが語られる（例えば、[3]）。すなわち、生物は、その進化過程において、生存・生殖に係るいわゆる「本能」が形成され、その中で、感知・行動システムや「感情」などが形成された。そのような中、生物から進化した人間（ホモサピエンス）においては、道具の利用などの「変異」が、個体間の情報交換を通じて社会全体に普及し、適者生存として機能してきたことが語られる。さらに、農業革命以降に形成された社会の中で、人々は言語による高度なコミュニケーションと「意識」の創造がなされ、それに伴い、ロジカルな思考による推論などを含む「知性」が発展するとともに、社会・技術のイノベーションが進展してきたとされる。その後、社会制度の変化に伴う産業革命により、現代に見られるような市場メカニズム等を通じた急速なイノベーションの時代が到来することになるが、その際、言語による高度なコミュニケーションが重要な役割を果たしてきたことは言うまでもない。特に、過去における情報通信・デジタル技術の進展は、人類によるイノベーション活動を大きく加速してきた。具体的には、印刷技術の進展は、既存の思想を広く多くの人々の知的活動の基盤として発展し、電話をはじめとする通信技術の進展はもちろんのこと、インターネットの登場は、人間の知的活動における情報収集・交換プロセスを抜本的に変革してきたと言える。

このような中、近年における AI（人工知能）技術の登場は、(1) に示した現代の社会経済システムにおけるイノベーションのプロセスに大きな影響を与えただけでなく、これまでイノベーションを起こしてきた人間の知能（知性）に AI はどこまで近づくのか（AI はイノベーションを起こすことができるのか）という論点も含めて、過去の人類・社会の発展において培われてきた人間によるイノベーションのプロセスの在り方の本質に対して与える影響についても、関心が高まっている。

² なお、需要者・市場は常に変化しており対象は常に変化していくことから、唯一の最適解が存在するものではないことに留意することが必要である。

3. 生成 AI システムの技術的な仕組みとその限界

(1) 対話型生成 AI (ChatGPT など) の技術的な仕組み

近年、爆発的に普及しつつある対話型生成 AI システムは、人間の知性の主たる表現手段である言語を扱う汎用性なシステム)、また、質問に対するスムーズな回答ぶりから、いわゆる AGI とされるような人間の知的能力に匹敵するような能力を持ち始めたとの指摘も一部にある。特に、過去の膨大な知識を学習することにより、一人の人間が有する知識量を圧倒的に超えるような知識能力を有するという特徴を有する。

しかしながら、その技術的な仕組みを見ると、その基盤となる大規模言語モデル (LLM) は、端的には、過去の膨大な文章に基づき次の単語を予測する機械にしか過ぎないと言える (例えば、[4])。具体的には、この対話型生成 AI システムは、技術的な観点からは、基本的に、以下の二つの仕組みから構成される (図 2 参照)³。いずれの仕組みも、人間による過去の知識と人間による回答パターンなどを学習することによって構成された機械にしか過ぎないことが分かる。

- 1) 過去の人類による膨大な (ネット上の) テキストデータ (文章) 等を集めることにより、確率論の観点から次の単語を予測するシステム (大規模言語モデル LLM) を構築。その際、深層学習等を行うことにより、人間が書くような「文章」を生成。
- 2) その LLM を利用しつつ、モデルの目的に合わせた適切な回答を行うことができるよう、人間のフィードバックによる強化学習 (RLHF) を含む各種の機械学習技術を適用。これにより、人間の回答パターンを学習させ、質問に対する、人間らしい「回答」を生成。

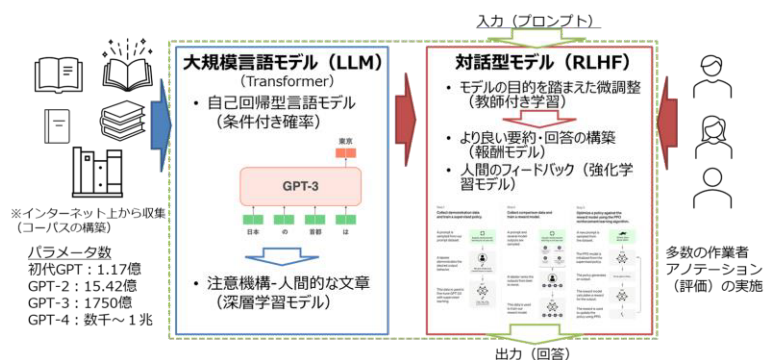


図 2: 対話型生成 AI システム (ChatGPT) の技術的な仕組み (出典) 市川[4]を踏まえ、筆者修正。

このように、ある意味、次の単語を予測する機械にしか過ぎないため、本質的に嘘 (Hallucination) を起こし得ることに加え、推論 (論理構成) や計算は不得意であることが指摘されている。ただし、データとモデルの大規模化により、「創発」 (予想外の能力の創出) が生じたとされ、ある種の知識データベース「的」な位置づけを有するに至ったとも解釈されている。

(2) 現在の対話型生成 AI の技術的限界

その際、この対話型の生成 AI システムは、その知識量は膨大であるものの、原則として、インターネット上に存在する人類の過去の知的活動の成果である大量の文章・知識の蓄積を学習した機械にしか過ぎない。このため、以下のような限界が内在する。

- その AI システムが学習した、過去の人類によって蓄積された知識・英知を超えるような回答は、原則として期待できない。すなわち、新たな推論などの導入によって、新しい知識を創出することは、基本的にはない。ただし、新たな「組み合わせ」は提示する可能性はある。
- 単語予測システムであることから、原則としては、もっとも確率の高い文章、すなわち、多数派の文章が生成される。さらに RLHF 等の設計にも依るもの、一定の「優等生的」な回答に偏り、異端ではあるものの革新的な回答は期待できない可能性がある。ただし、プロンプトを通じた対話の工夫により、必ずしも多数派ではない回答を探し当てることは可能である。

このような限界が、第二章で述べたイノベーションのプロセスにどのような影響を与えるかについて、それぞれ、以下、第四章、第五章で議論する。

4. 生成 AI との対話によるイノベーションプロセスの合理化

(1) 生成 AI のイノベーションプロセスへの適用可能性

それでは、この現在の対話型生成 AI システムは、イノベーションのプロセスに対して、どのような影響を与えるのであろうか。以下、第二章 (1) で示した A 段階、B 段階それぞれについて議論する。

³ なお、ChatGPT 等の LLM に関しては、その詳細な仕組みは公開されておらず、ブラックボックスとなっており、このため、現在多くの AI 研究者が、ChatGPT を試行し、その回答結果を踏まえて、そのブラックボックスの中身を推定し、また、その性能を評価しようとしているところである。

＜アイデア（新結合）の創出段階（A 段階）＞

まず、イノベーションのプロセスにおけるアイデア創出（A 段階）において、生成 AI は大きな役割を果たすことが期待・想定されている。すなわち、上述の通り、対話型生成 AI システムは、原則としては、新たな知識創造活動を行うものではないものの、一人の人間がとても扱うことができないような膨大な既存知識を学習している。このため、これまで多数の関係者・知見者と相談・コミュニケーションしながら進めていたアイデア生成のプロセスについて、対話型生成 AI と対話することによってかなりの部分を代替することが可能になると考えられる。さらに、プロンプトを工夫することにより、利用者の期待に応じた回答に近づけていくことも可能である。

また、上述のように原則として新たな知識創造活動を行うものではないものの、実際に、生成 AI の回答は、創造的な場合もあるとの報告もある。これは、創造性とは複数の知識の組合せによって生じるものであることを踏まえると、例えば、文献 X に記載している内容と文献 Y に記載している内容を、これまでになかったような新たな組合せ（アイデア）を提示することによって、一定の創造性を発揮し、新たな視点を提供するような場合などが考えられる。

＜イノベーション（新結合）の遂行段階（B 段階）＞

一方、イノベーションの遂行段階（B 段階）においては、少なくとも現状の生成 AI は、原則としてほぼ対応していないと考えられる。すなわち、現状の生成 AI は、与えられた質問に対して一定のパターンの回答を作成する機能のみであり、人間のように自ら目標を決定し、その目標の達成に向けて複雑な思考プロセスを行う機能、また、それを現実社会において関係者を巻き込んで目標達成に向けて実現していく機能は有さない。このため、このような A 段階で得られたアイデア・構想を、現実社会において実現・遂行可能で社会ニーズ（目的）に照らして最適なアイデアなものとして具体化し、確立・実現（遂行）していくプロセスに関しては、引き続き、人間が中心に担うこととなる。

ただし、デジタルだけで完結する一部の分野では、将来的には、生成 AI 以外のデジタル・AI 技術なども組み合わせる利用することによって、これらのプロセス全体を自動化する仕組みを構築することも、想像可能である。ただし、その場合には倫理面等についても考慮することが必要になる。

（2）イノベーションプロセスにおける人間と生成 AI との対話

ただし、これらの A 段階と B 段階は、両方とも試行錯誤を通じて互いに関連しあいながら進めるものである。このため、今回の生成 AI の登場により、引き続き人間が中心となって、特に A 段階を中心に必要に応じて生成 AI と対話をしながら試行錯誤を行うことにより、イノベーション創出に係るプロセス全体の効率化を図ることが可能になるものと考えられる（図 3 参照）。

その際、いわゆるプロンプトエンジニアリングと言われるような、質問内容を工夫しながら、より適切な回答を求めていくプロセスが重要な手法となる一方、今後の対話型の生成 AI においては、いわゆる人間・機械チームング（Human-Machine Teaming : HMT）と言われるような、人間との関係を踏まえた研究開発が重要になるものと考えられる。

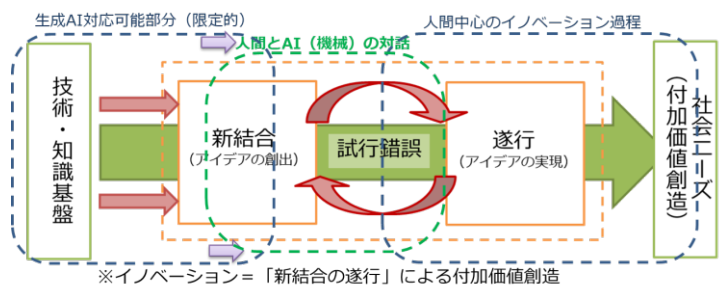


図3：人間とAIとの対話（HMT）によるイノベーション創出の促進
（出典）市川[4]を踏まえ、筆者修正

5. 生成 AI がイノベーションにもたらす長期的な懸念と人間の役割（考察）

（1）「再帰の呪い」がイノベーションに与える長期的な影響

一方、上述の通り、現在の生成 AI は、基本的には次の単語の予測システムであり、原則として、過去の文章の蓄積から最も確率の高い文章を生成するものである。それゆえに、生成 AI の利用は、「再帰の呪い」の問題として、過去にわたって人間により蓄積されてきたイノベーションの基盤である技術・知識基盤を損なう可能性が指摘されている。ここで、「再帰の呪い」とは、生成 AI から生成・出力された文章等を、再度生成 AI に入力し学習させるというプロセス（再帰）を繰り返した場合、文章としての確率の高いもの（大きく言えば、多数派の文章）のみが残り、基盤モデルとしての多様性が全く失われてしまうという現象を指す（[6]参照）。このことは、仮に、人間が関与することなく現行の生成 AI から出力された文章等が多量にネット上に公開された場合、ネット上の文章は汚染され、そのネット上の文章を学習する LLM は崩壊しかねないこと、ひいては、過去から人間によって生成・蓄積され、人類

のイノベーションの基盤となってきた膨大で多様な文章空間である技術・知識基盤が失われることになりかねないこと意味する。これは、逆に言えば、少なくとも、現在の生成 AI をベースに考える限り、これまでの人間中心のイノベーションを進めていくためには、引き続き人間の関与が不可欠であることを意味する。

また、再帰の呪いはあくまでも長期的な懸念であるとしても、少なくとも、現行の生成 AI の確率の高い（多数派）の回答を行うという特徴は、そのシステムの調整にも依るものの、イノベーションを起こす要因である多様性を減ずる可能性があるとともに、通常では思いつかないような組み合わせである Serendipity を十分起こせない可能性もあることに留意することが必要である。実際に、生成 AI で生成された SF 小説などは、非常に平凡（ワンパターン）であるとの指摘もなされている。

（2）何故、人間はイノベーションを起こすことができるのか（再帰の呪いからの解放）

それでは、このような生成 AI での限界に対して、これまで、人類・社会においては、なぜ、多様な言語空間を形成することが可能であったのであろうか。これらについては、そもそも人間は、個人それぞれの多様な見方が存在し、それらが言語化され公表されているという大前提に加えて、例えば、以下のような要因が考えられる。

- ・ ①現実社会の言語化：まずは、人間においては、その発展プロセスにおいて、現実社会の事実関係を感じ、それを言語化・伝達する能力・本能が備わっていること。これに対し、少なくとも現時点の AI においては、現実社会における事実関係をデータとして取得することは可能であるものの、それを言語化し、言語としての新たな知識として蓄積することが可能な状況にはなっていない。
- ・ ②推論を通じた論理構築能力：その上で、人間においては、今後の方向などについて、主観的でかつ多様ではあるものの、言語に基づき推論するなどの論理を構築する能力があること。これに対し、現時点の生成 AI では、言語に基づく推論を行い予測する能力に関して限界がある。
- ・ ③生存本能を踏まえた目的の設定：さらに、その前提として、人間は、個々に応じて生存本能に基づき目指すべき目的・目標を設定しうる。これに対し、AI システムは、原則として、人間の設定した目標に対して、せいぜい最適解を与えることしかできず、特に上記推論が困難であれば、その解も出せないこと。

ただし、AI 技術がさらに進展すれば、これらの人間の機能の一部を AI に任せることが可能になる時代も来る可能性がある。さらに、上記の議論は、これまでの言語による人間中心のイノベーションを想定しているが、観念的にはデータのみによる計算論に基づくイノベーションの可能性も考え得る。ただし、そのような AI の社会での利用に関しては、現在生成 AI において世界的に議論されている以上に、AI と人間社会の在り方（倫理問題等）が問われることになると考えられる。

参考文献

- [1] シュムペーター「経済発展の理論 企業利潤・資本・信用・利子および景気の回転に関する一研究（上・下）」、岩波書店（1977）
- [2] シドニー G. ウィンター、リチャード R. ネルソン「経済変動の進化理論」、慶應義塾大学出版会（2007）
- [3] マット・リドレー「人類とイノベーション:世界は「自由」と「失敗」で進化する」NewsPicks パブリッシング（2021）
- [4] 岡之原大輔「大規模言語モデルは新たな知能か ChatGPT が変えた世界」、岩波科学ライブラリー 319、岩波書店、（2023）
- [5] 市川類「第四次 AI ブーム（ChatGPT）による世界の AI ガバナンス制度の進化 ～ChatGPT 型 AI システムの社会的リスクと世界の AI 規制・ガバナンス政策の動向～」、WP#23-02 一橋大学イノベーション研究センター（2023/05/15）
- [6] Ilya Shumailov, Zakhar Shumaylov, Yiren Zhao, Yarin Gal, Nicolas Papernot, Ross Anderson, “The Curse of Recursion: Training on Generated Data Makes Models Forget, arXiv:2305.17493, (27 May 2023)