

生成AI活用のための アジャイル・ガバナンス

生成AIのイノベーションを素早く業務に活用し、生産性や付加価値を向上させることは経営上極めて重要である。一方、適切なコントロールへの社会的な要請もある。生成AIの活用推進を止めず、かつ安全に活用できるようなガバナンスモデルとしてアジャイル・ガバナンスが注目される。

生成AIとAI事業者ガイドラインの 公開

自然言語処理や画像生成などの分野で生成AIが急速に発展している。金融機関においても生成AIを活用した業務効率化やお客様への更なる価値提供を目指し、ビジネスへの活用を模索するケースが急増している。

一方で、AI技術の利用範囲や利用者の拡大に伴いリスクの増大も懸念されている。こうした問題意識について、政府は2019年3月、来るべきSociety 5.0に備えるため、「人間中心のAI社会原則」（統合イノベーション戦略推進会議）を策定し、AIが適正に利用されるための指針とビジョンを提示した。

その後、生成AIの出現により知的財産権の侵害、偽情報・誤情報の生成発信など、これまでにないリスクが大きな課題として表面化したことから、2024年4月19日には総務省と経済産業省は「AI事業者ガイドライン（第1.0版）」を公開するなど、さらに一歩踏み込み、生成AIの安心安全な活用を促進するためのガバナンスに関する議論を活発化させている。

生成AI活用を模索する金融機関にとっても生成AIガバナンスの構築は今後、検討していかなければならない必須のテーマといっても過言ではない。

生成AIガバナンスの必要性

生成AI技術の進歩やビジネスでの活用があまりにも急速に進むため、ガバナンス面の整備の遅れを懸念する声も出始めている。例えば、米国の非営利団体がOpenAI社のGPT-4よりも強力なAIシステムのトレー

ニングを少なくとも6か月停止するよう求めたオープンレター¹⁾は、著名なAI研究者や技術者、経営者が署名したことで話題になった。これは、AIシステムの開発スピードに社会や規制当局の対応が間に合わないことを懸念したものである。

日本においても、イラストレーターや漫画家などになる団体による画像生成AIの適切な使用や法整備などを求める動きも見られたが、これもガバナンスの不備を表すものだろう。

ガバナンスの整備の遅れは企業活動にも影響が生じており、例えば、イラスト作成アプリに、生成AIによる「お手本」を生成する機能を搭載したところ、ユーザーからの反対意見などが相次ぎ、リリース翌日に機能を撤回する事態も発生した。

生成AIにおいて、従来にはない特有のリスクをコントロールするためのガバナンスが考えられる。従来のシステムでは、目指すべき目的も比較的設定しやすく、例えば、セキュリティの観点から情報の機密性や完全性を確保する、システムの可用性を高めるといったゴール設定が明確化されていた。また、動作が比較的予測可能であるため、設計の段階で相対的にリスク分析を行いやすく、それに伴う対策も可能であった。さらに、技術に対する社会の受容度合いも高かったといえるだろう。

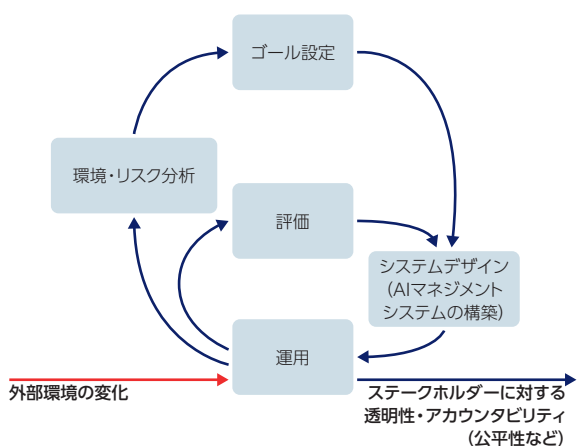
しかし、生成AIを活用したシステムにおいては、「人間中心、公平性、教育・リテラシー」など²⁾、より人間の本質的な価値への配慮を見据えたゴール設定も求められる。

また、生成AIは設計段階で事前に動作を完全に予測することが困難なシステムや、動作のたびに結果が変わりうるシステムになることも考えられる。さらに、社会

NOTE

- 1) Pause Giant AI Experiments: An Open Letter (<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>)
- 2) 「AI事業者ガイドライン(第1.0版)」における指針の例。
- 3) 「AI事業者ガイドライン(第1.0版)」においても、アジャイル・ガバナンスの実践が重要である旨が記載されている。

図表 アジャイル・ガバナンスの基本的なモデル



(出所) AI事業者ガイドライン 第1.0版

の受容度合いも日々刻々と変化しており、ある時点で受け入れられていたAIシステムが、ある時点では受け入れられない、またはその逆となることも短期間で起こりうる。このため、従来のシステムと異なる観点も考慮したガバナンス構築が必要となる。

このようにゴール設定が明確でなく、加えて日々刻々と変化しうるAIシステムのリスクに対応するためには、「AI事業者ガイドライン」の中でもその重要性が言及されている通り「環境・リスク分析」、「ゴール設定」、「システムデザイン」、「運用」、「評価」といったサイクルを回転させていくことが求められる。これを実践するためには、“アジャイル・ガバナンス”のアプローチ³⁾が効果的である(図表)。

アジャイル・ガバナンスの要諦

アジャイル・ガバナンスを実践するには、まず生成AIがもたらす便益とリスクを分析し、その結果を経営

層も含めて共有し適時に理解を更新することが重要である。加えて、社会的受容度合いを意識しマルチステークホルダーの意見も取り入れながら、人間中心の価値や公平性確保といったガイドラインの指針などに基づいたゴールを設定する。

例えば、金融機関では生成AIを活用したサービスの提供において多様性・包摂性、公平性などが担保されることはゴール設定における重要な論点の一つになる。

その後、設定したゴールを達成するために「システムデザイン(AIマネジメントシステムの構築)」を行ったうえで「運用」する。その際には、AIマネジメントシステムを適切に運営するためのリテラシー向上を図ることや、主体間・部門間での協力を行える事前準備などもポイントとなる。必要に応じて運用状況を適切に開示できるような説明可能性の確保も肝要である。そして、多くのステークホルダーの意見を求めながら、ガバナンスの機能を定期的に評価し、継続的な改善につなげることが重要である。

つまり、アジャイル・ガバナンスの要諦は、先述した環境・リスクの分析、ゴールの設定、システムのデザイン、運用、評価のサイクルを、マルチステークホルダーとともに継続的かつ高速に実行することにある。とりわけ多様なステークホルダーとの継続的な関係を構築していくことがポイントとなる。

Writer's Profile



金子 洋平 Yohei Kaneko
金融デジタルビジネスリサーチ部
シニアリサーチャー
専門はリテール金融、金融DX
focus@nri.co.jp