

未知の AI リスクに備える

—エネルギー・軍事・金融領域におけるリスク管理の学びを活用せよ—

CDO ユニット チーフデジタルオフィサー 大野 有生（専門分野：AI・デジタル活用）

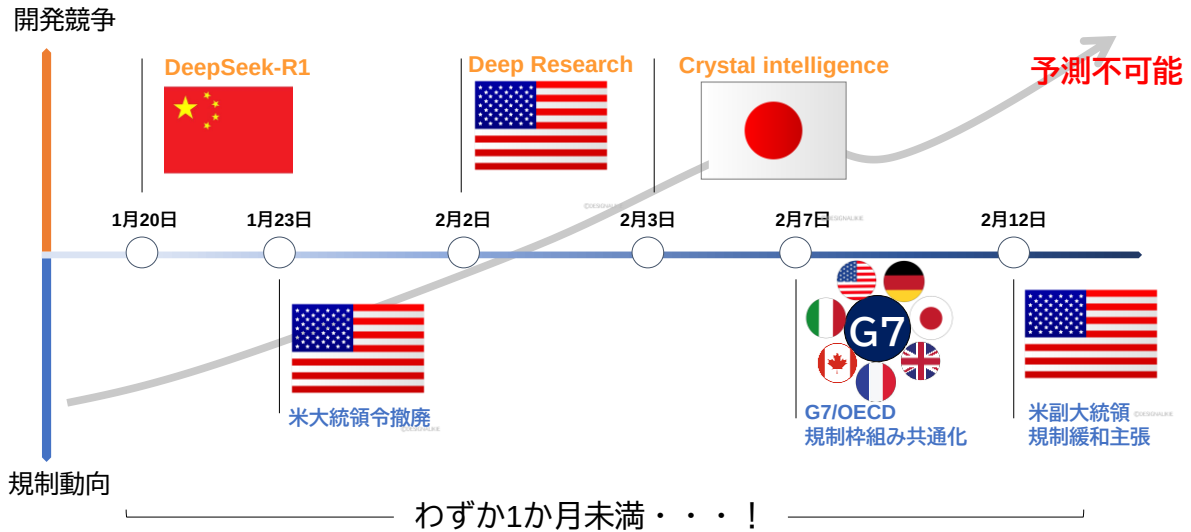
1. AI の急速な進化と最新規制動向

AI 技術は、100 年に一度と言われるパラダイムシフトの最中にある。ここ 1 カ月の情勢を見ても、その変化の激しさは明白である。米国では 2025 年 1 月 23 日に、トランプ政権がバイデン前大統領の AI 規制大統領令を撤回し、革新的な研究および利用の促進を強調した^[1]。わずか 4 日後、中国の AI 企業「DeepSeek」による革新的な低コスト高性能 AI モデルの発表（1 月 20 日）の影響により、米国市場に激震が走り、NVIDIA の株価は一時 17% 下落し、時価総額で 5,900 億ドル（約 91 兆円）が消失した^[2]。一方、2 月 2 日には米国 OpenAI は Deep Research を発表^[3]、そして、翌 3 日にはソフトバンクグループと共同で「SB OpenAI Japan」を設立し、企業専用 AI「Cristal intelligence」を超知能へと進化させ、経営判断やシステム開発の効率化を目指すと発表した^[4]。

一方、AI の急速な発展に伴い、そのリスクも拡大している。2023 年 5 月の G7 で合意された「広島 AI プロセス」を皮切りに、各先進国政府や国際機関は、法的・倫理的な枠組みの整備を進め、AI 技術の安全な開発および利用に向けたルール作りを加速している。2025 年 2 月 7 日には、G7 各国が生成 AI の事業者に対して安全性に関する情報提供を求める運用を開始した。OECD が中心となり、40 か国以上の企業や研究機関が参加する中、米マイクロソフトやグーグル、日本勢の NTT、NEC、KDDI などがその意思を表明している^[5]。しかし、米国はこれに批准せず、2025 年 2 月 12 日に、バンス副大統領は、AI ACTION SUMMIT にて、世界のリーダーたちに対し、AI に関する規制をけん制した^[6]。将来的に規制強化の動きが生じれば、政治的摩擦が避けられない可能性がある。

驚くべきは、これら全ての衝撃的ニュースの数々は 2025 年に入りわずか 1 カ月未滿で起こった出来事であることだ（図表 1）。AI をめぐるグローバルレベルでの攻防は、加熱の一途を辿り、リスクは膨らむばかりだ。

図表 1 2025 年の AI に関する出来事（筆者作成）



2. AI 規制の現状と未知リスクへの備え

現行の AI 規制に関する議論は、主にアルゴリズムの不透明性やデータバイアスなどの技術的側面に焦点が当たっている。たとえば、経済産業省が発行する AI 事業者ガイドライン^[7]では、倫理、品質、セキュリティの 3 つの観点からシステムの安全性、説明可能性、信頼性を高める対策が強調されている。しかし、AI の影響は技術領域に留まらず、経済、社会、政治、環境といった多様な分野に波及しているのが現実である。

さらに具体的には、現代社会が直面する課題として「説明責任の曖昧化」や「社会的偏見に基づく倫理観の歪み」、さらには「フェイクニュースの拡散による地政学的リスク」など、従来の情報技術対策では対応が難しい新たなリスクが顕在化している。これらのリスクは、未知の大惨事を招く可能性すら孕んでおり、単一の技術的対策だけでは十分でないことが明らかである。

このような背景から、未知のリスクに備えるためには、これまで他分野で蓄積されたリスク管理の知見を応用することが求められる。他分野での成功事例を踏まえれば、AI リスクにも同様の包括的アプローチが必要であると考えられるためである。

そこで本稿では、リスク管理が経営の根幹であると長らく認識され、その手法が確立されているエネルギー・軍事・金融という 3 つの分野に注目する。これらの業態は、歴史的にも厳しい環境下でのリスク対策を重ね、実績を上げてきたことから、AI リスク管理への応用可能性についても示唆に富んでいる。具体的には、各分野で確立されたリスクマネジメント手法を取り上げ、これらの知見をどのように AI リスク管理に応用できるかを考察する。最終的には、異なる領域から得られた知見を統合し、「AI ガバナンス」を強化することで、未知の AI リスクに効果的に対処する具体的なアプローチを提示することを目的とする。

3. 他領域のリスク管理から学ぶ

(1) エネルギー業界

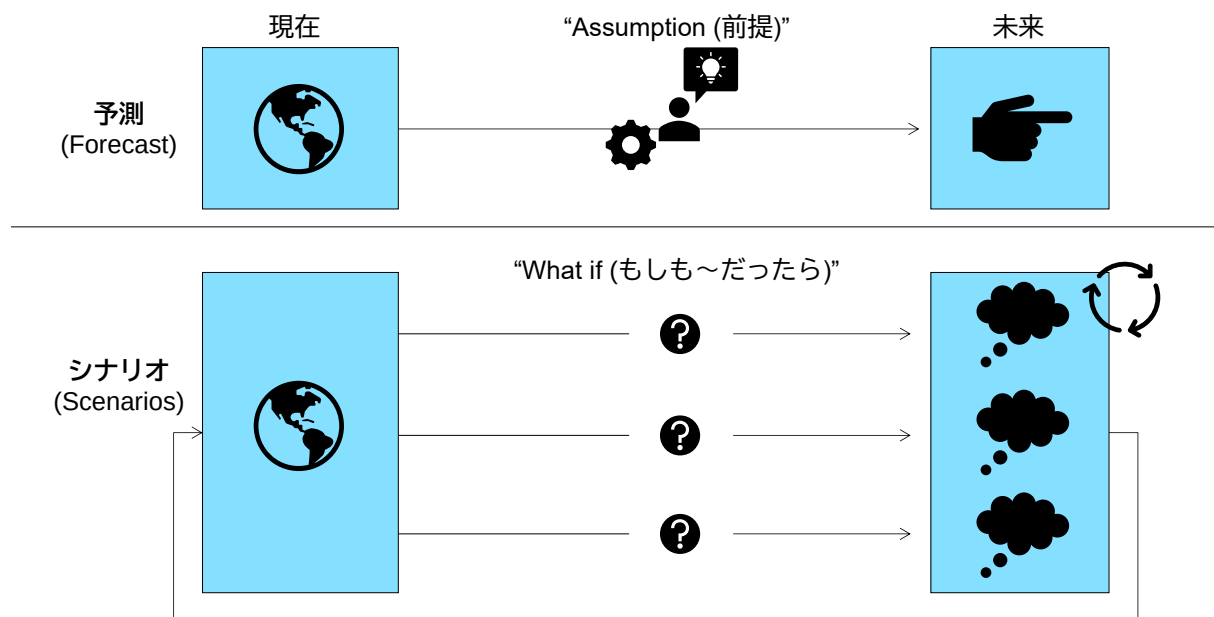
エネルギー業界は、環境変化や市場リスクの影響を強く受けるため、リスク管理が経営戦略の中心に据えられている。特に、「シナリオプランニング」や「ルール責任転換モデル」を適用することにより、激動の時代を生きながらえてきた。

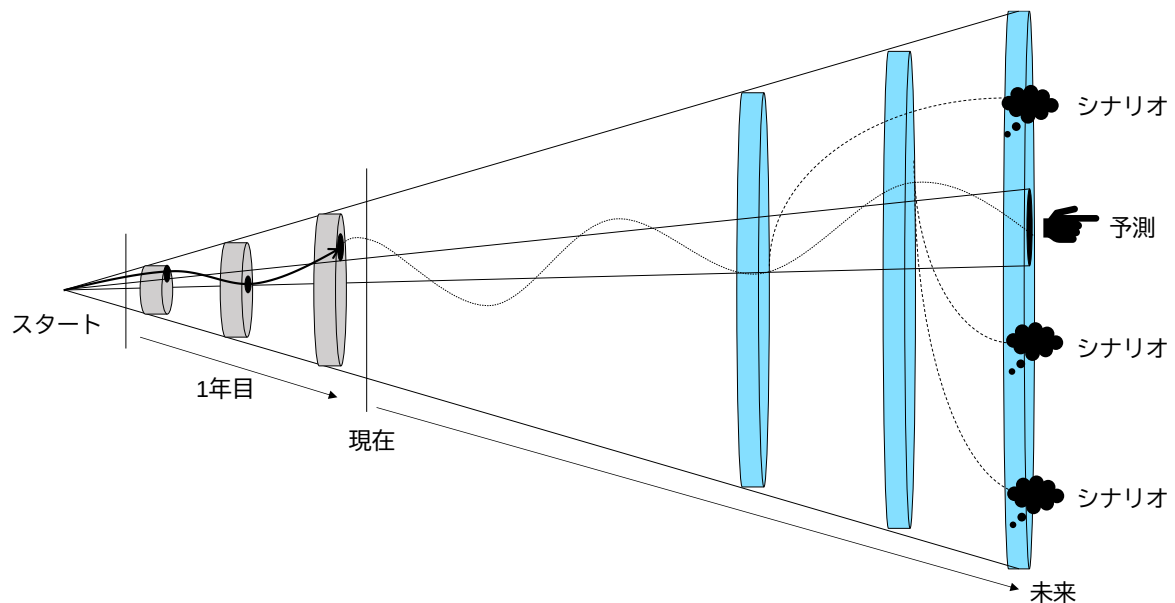
① シナリオプランニング^[8]

ロイヤル・ダッチ・シェルは、50 年以上前の 1972 年に「シナリオプランニング」、すなわち、将来起こりうる複数のシナリオを描き、事前に対策を立てる手法を確立し、以後、3 年に一度、シナリオを更新し続けている。

シナリオプランニングとは、未来の「予測」とは異なり、複数の「ストーリー」として提示する手法であり、経営者が環境の不確実性を認識し、適応策を講じるためのツールである(図表 2)。

図表 2：予測 vs シナリオ





(出典) Three Decades of Scenario Planning in Shell^[8], Integrating organizational networks, weak signals, strategic radars and scenario planning^[9]をもとに筆者作成

シェルは、1970年代からエネルギー市場中心のシナリオを展開し、1980年代以降は経済、政治、社会の要素を統合したグローバルシナリオへと進化させた。1990年代には外部ステークホルダーを巻き込み、環境変化や社会トレンドを重視する姿勢を取り、2000年代にはグローバルシナリオと事業課題に焦点を当てたシナリオが統合され、戦略的意思決定の基盤となった。近年は地球温暖化やテロリズムといった予測困難な課題に対しても重要な洞察を提供することに成功している。さらに、シナリオは未来の選択肢を明示し、これにより、地域別や事業別の詳細なシナリオが各プロジェクトの意思決定に活用される仕組みが整えられている。

シナリオプランニングは未来の不確実性を包括的に捉え、経営判断の質を向上させるための有力なフレームワークである。シナリオは正確な予測を示すものではなく、経営者が洞察を深めるための枠組みとして位置付けられる。

② ルール責任転換モデル～個人を責めず、ルールを責める^[10] (Oil & Gas IQ Editor, 2019)

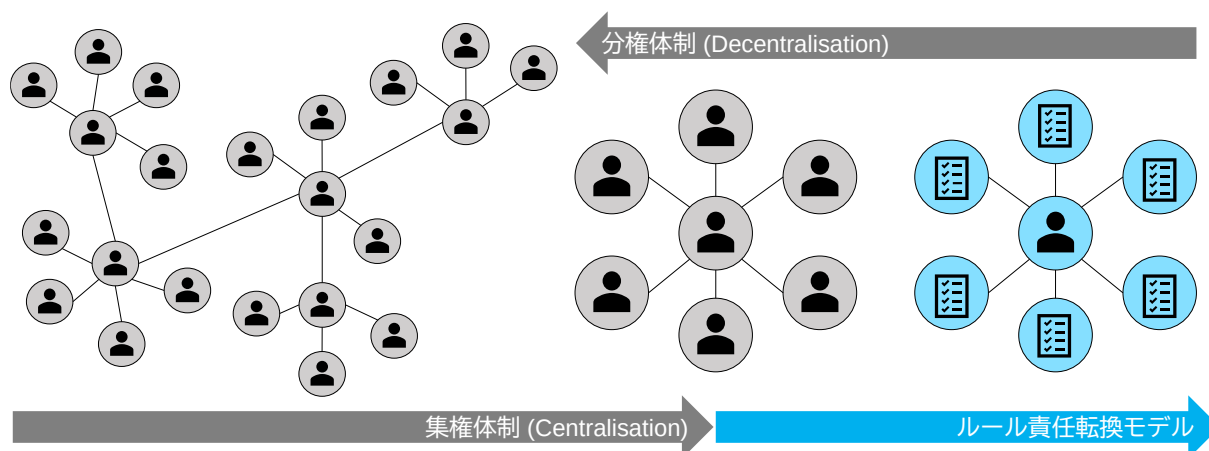
2010年4月20日、英エネルギーメジャーBP(British Petroleum)が運営するMacondo油田においてDeepwater Horizon掘削リグが爆発・火災し沈没した事故は、米国史上最大の原油流出事故となった。87日間で約490万バレルの原油が流出し、BPは市場価値の大幅喪失とともに、事故処理費用、罰金、賠償金として564億ドル(2010年レートで約5.2兆円)を支払う結果となった。事故の直接的要因は、セメントバリアの失敗、安全テストの誤解、および検知と対応の遅れである。

BPは1990年代以降、組織改革によりリスク管理を分権化し、現場レベルでの迅速な意思決定を促進した。しかし、この改革は全社的なリスク監視の弱体化を招き、分権化によるガバナンスの不備が深層的なリスク要因として指摘される。

事故後、BPは中央集権型のリスク管理体制への回帰、安全基準の再評価、および新たな規制の導入を進めた。重要であるのは、BPでは「(インシデントにつながる)ルール違反が発生するのは、個人のモラルよりも、そもそもルールが不明確・不適切だから」という考え方を徹底したことだ。ルール違反の理由を「ルールが不明確である

か、ルールが個人のアサインメントに対して適切でないからである」とし、ルールを責め、リーダー層が改善する義務を持つ。それをリーダー層の評価につなげることで、組織全体としてリスクを軽減する仕組みを構築している。

図表 3：ルール責任転換モデル（筆者作成）



図表 3 が示す通り、企業規模が大きくなりビジネスモデルが複雑化するほど権限移譲が進み、分権体制が進むが、個々人が見えている範囲での個別最適の意思決定を行うために、システムや会社全体を巻き込む大規模事故が発生する確率は高くなる。少しでも大規模事故に関連すると思われる事象を発見した際には、意思決定は上位に委ねるべきである。さらに、意思決定を下す上位職責者を責めることなく、責任は意思決定基準が記載されるルールそのものにする仕組みを構築するべきだ。

(2) 軍事業界

軍事業界では、外部環境の変化に適応し、早期警戒システムを構築するために「戦略レーダー」や「微弱信号の監視」を活用し、予測困難な状況に備えている^[9]。

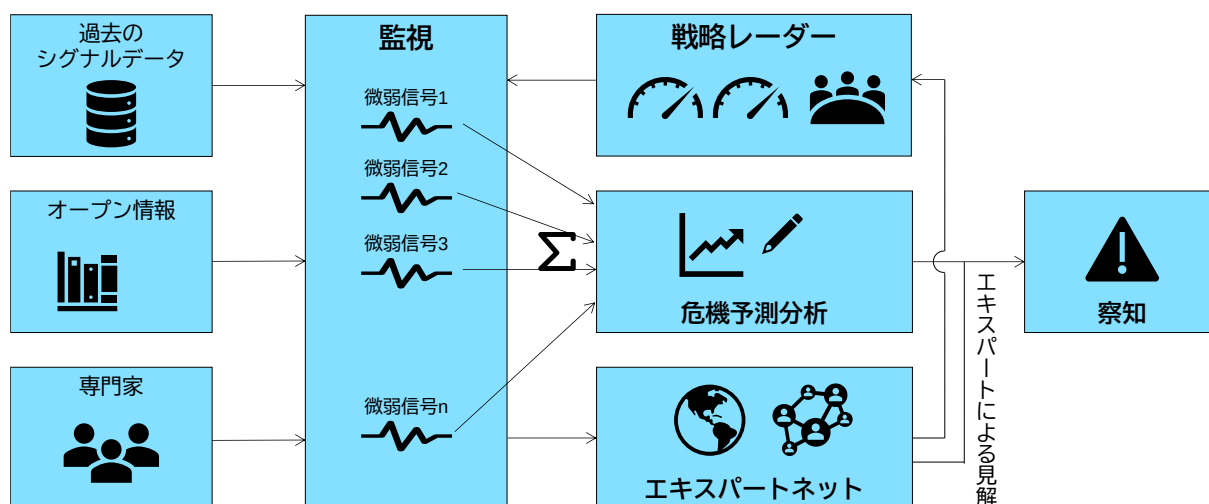
□ 戦略レーダーを活用した微弱信号の検出 (Strategic Radar and Weak Signals)

軍事組織では、外部環境の変化に迅速に適応するために「**戦略レーダー (Strategic Radar)**」を活用している。戦略レーダーとは、組織の外部ネットワークを活用し、新たな脅威や機会の兆候を早期に察知するシステムのことである。

さらに、戦略レーダーの監視対象として重要なのは、重大な危機の前兆となる「**微弱信号 (Weak Signals)**」である。微弱信号とは初見では無関係または背景ノイズのように見えるものの、適切に解釈されれば重要な意味を持つ情報である。このため、企業が微弱信号を見逃さないためには、広い周辺視野が必要である。

たとえば、米国防総省の防衛物流局 (DLA : Defense Logistics Agency) は、戦略レーダーを実装し、エネルギー、技術影響、地政学的リスクといったメカテゴリーで微弱信号を常時監視していた。冷戦時代の情報戦では、敵国の兵站や通信の微妙な変化を微弱信号として捉え、RFID 技術の衰退やエネルギー価格上昇などの傾向を早期に検出し、戦略調整を実施している。このように、些細な変化をいち早く捉えることが、先手を打つリスク管理には欠かせない。

図表 4：微弱信号監視と危機予測分析



(出典) Integrating organizational networks, weak signals, strategic radars and scenario planning をもとに筆者作成

(3) 金融業界

金融業界は、リスク管理手法を学術的理論から事業に適用した初めてのドメインである。従って、長年にわたりリスク管理手法が発展してきた。

□ エスカレーション奨励文化

元来、金融業界ではリスク測定のツールが不足していたため、適切なリスク管理が難しかった。一方、学術界においては、1940年代からリスクとリターンの関係を研究し始め、1970年代から金融理論—ブラック＝ショールズ方程式や資本資産価格モデル（CAPM）—が登場し、金融市場のリスクを定量的に測定する技術が発展した。さらに、モンテカルロシミュレーションや Value at Risk（VaR）などの手法により、リスクの可視化とヘッジが可能になった^[11]。

しかし、いくら工学的に発達した高度なリスク管理モデルを駆使してきた金融業界も、ブラックスワン(想定外の出来事や大きな衝撃をもたらす現象)^{脚注1}には太刀打ちできなかった。2007～2008年に、金融システム全体が崩壊の危機に陥ったのだ。原因は、サブプライムローンや信用リスクの過小評価によるものであり、多くの金融機関が数十億ドル(数兆円以上)の損失を出した。

しかし、ゴールドマン・サックスは、他社に比べてサブプライム関連の評価損失額がけた違いに小さく、2007～2008年を通じて純利益を確保している。2008年においては、モルガンスタンレー(米国)・シティグループ(米国)・UBS(スイス)・ドイツ銀行(ドイツ)の4社純利益平均が約129億ドル(2008年換算レートで1.2兆円

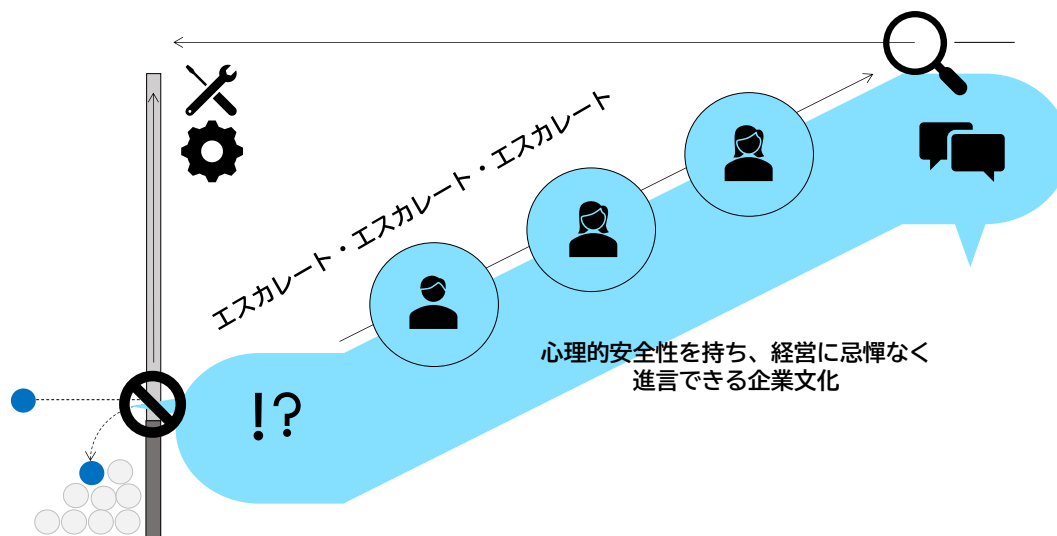
¹ ブラックスワンは、想定外の出来事や大きな衝撃をもたらす現象を指す言葉で、経済・投資・社会情勢などのさまざまな分野で使用される。この言葉の由来は、黒い白鳥が発見されたことによる衝撃から来ている。ブラックスワンに対する備えと対策は重要であり、事前の情報収集やリスク管理が不可欠となる。

近く)の赤字であったのに対し、ゴールドマン・サックスは 23.2 億ドル(同 2,200 億円)の黒字で着地している[12]-[16]。

ゴールドマン・サックスが損失を最小化できた要因は、他の投資銀行に比べて早い段階でリスクに備えていたアジリティ(敏捷性)と、モデルによる定量分析だけでなく、経営陣の判断による定性的分析も併用するアプローチをとっていたからだと言われている^{[17],[18]}。

ゴールドマン・サックスは「エスカレート・エスカレート・エスカレート（報告・相談を重ねる）」というルールを徹底していた(図表 5)。

図表 5：エスカレーション奨励文化（筆者作成）



リスク管理部門が心理的安全性を持ち、経営陣に忌憚なく進言できる企業文化により、トップダウンでのリスク議論が活発であった。これにより、硬直的でリスク兆候の共有が遅れた競合他社との差異化要因となり、危機を乗り切ることができた。

つまり、高度化された金融工学によるリスク管理手法でも未曾有の危機は回避できず、企業文化(ソフト面)でのリスク管理が重要であることが伺える。

4. 他業界のリスク管理手法の AI リスクへの適用可能性検証

(1) ケンブリッジ大学リスク研究センターのリスク分類

ケンブリッジ大学リスク研究センターは Cambridge Taxonomy of Business Risks（ケンブリッジ・ビジネスリスク分類）^[19]において、企業が直面するリスクを体系化している(図表 6)。

図表 6 : ケンブリッジ・ビジネスリスク分類

リスククラス（主要カテゴリ）	説明
財務リスク（Financial）	経済の変動、金融市場の混乱、企業の信用リスクなど
地政学リスク（Geopolitical）	戦争、政権交代、政策変更、規制リスク、政治的不安定
技術リスク（Technology）	サイバー攻撃、インフラ故障、新技術の影響
環境リスク（Environmental）	自然災害、気候変動、資源不足
社会リスク（Social）	人口動態、消費者行動、健康トレンド、疫病リスク
ガバナンスリスク（Governance）	経営の失敗、コンプライアンス違反、訴訟リスク

(出典) University of Cambridge. (2019). Cambridge Taxonomy of Business Risks.
Cambridge: Cambridge Centre for Risk Studies.^[19]をもとに筆者作成

(2) エネルギー・軍事・金融業界の知見と AI ガバナンスの適合性

先述のエネルギー・軍事・金融業界で培われた知見を、フレームワークに適合させたのが図表 7 である。

図表 7 : 他業界のリスク管理手法とリスク分類の適合性（筆者作成）

他業界のリスク管理手法		ケンブリッジ・ビジネスリスク分類						AI リスク への応用
業界	管理手法	財務	地政学	技術	環境	社会	ガバナンス	
エネルギー	シナリオプランニング	✓	✓	✓	✓	✓	✓	未知への備え
	ルール責任転換モデル						✓	社内統制 (ガバナンス)
軍事	戦略レーダーと 微弱信号の検知		✓	✓	✓	✓	✓	危機の前兆を察知
金融	エスカレーション 奨励文化						✓	社内統制 (ガバナンス)

3つの業界で使われている全てのリスク管理手法は、「ガバナンスリスク」をカバーしている。

ケンブリッジ大学の研究では、企業が実際に報告しているリスクと、企業崩壊の原因となる実際のリスクとの間に大きなギャップがあることも指摘されている。特に「ガバナンスリスク」は過小評価されがちであるにもかかわらず、実際には企

業崩壊の最大要因となりうるリスクだという。ガバナンスが欠如すると、問題が発生した際に組織が適切に対応できず、被害を拡大させてしまうからである。

以上から、「ガバナンスリスク」をカバーする 4 つの管理手法は AI リスクを管理するためにも極めて有効であり、適用するに値することを主張できる。

5. AI リスク管理を強化するための具体的アプローチ

(1) 各業界の管理手法の AI リスク管理への適用

① シナリオプランニング（エネルギー業界の知見）

シェルでは、未来のリスクを想定する「シナリオプランニング」を長年活用してきた。これにより、**不確実性の高い未来のリスクに対して事前に対策を立てることができる。**

AI にも想定外の動作や社会的影響が生じる可能性があるため、複数のシナリオを想定し、事前に備えるシナリオプランニングの導入は極めて有効だ。「規制強化」「社会的反発」「技術の急激な進化」など AI 特有のリスクシナリオを幅広く描き、対応策を準備することが求められる。

② ルール責任転換モデル（エネルギー業界の知見）

BP では個人を責めるのではなく、システムやルールを見直し、改善することにより、組織全体としてリスクを軽減する仕組みを構築している。

AI リスクにおいても、異常な動作や問題が発生した際に、「担当者の過失」として終わらせるのではなく、**ルールそのものが適正かを問い直す文化**を根付かせることが重要である。

③ 戦略レーダーと微弱信号検知（軍事業界の知見）

軍事業界では、**戦略レーダーを活用し、微弱信号を捉えて早期にリスク対策を行う手法**が確立されている。米国防総省の DLA は、物流の変化や新技術の台頭をレーダーで常時監視し、問題が発生する前に対応している。

AI においても、戦略レーダーを導入することで、「**新たな規制の兆し」「社会的な反発」「技術的な欠陥**」といった**微弱信号をモニタリングし、リスクの顕在化を防ぐことが極めて重要である。**たとえば、AI モデルにおいて同じ内容の質問に対して、その回答が微妙に相違する程度を検知することで、AI モデルの品質・信頼性を担保し、大きな社会的問題を未然に防ぐことが可能になる。

④ エスカレーション奨励文化（金融業界の知見）

金融業界、特にゴールドマン・サックスでは「エスカレーション奨励文化」を組織に組み込むことにより、**バイアスを排除した定性的なリスク管理を行うことができる。**

AI リスク管理においても、技術者や現場担当者がリスクを察知した際に、心理的安全性を持ち速やかに上層部へ報告できるエスカレーション奨励文化を整えることで、組織のガバナンスを強化できる。たとえば、AI の判断がバイアスを含んでいると気づいた際に、それを無視せず速やかに報告できる体制を作ることが重要である。

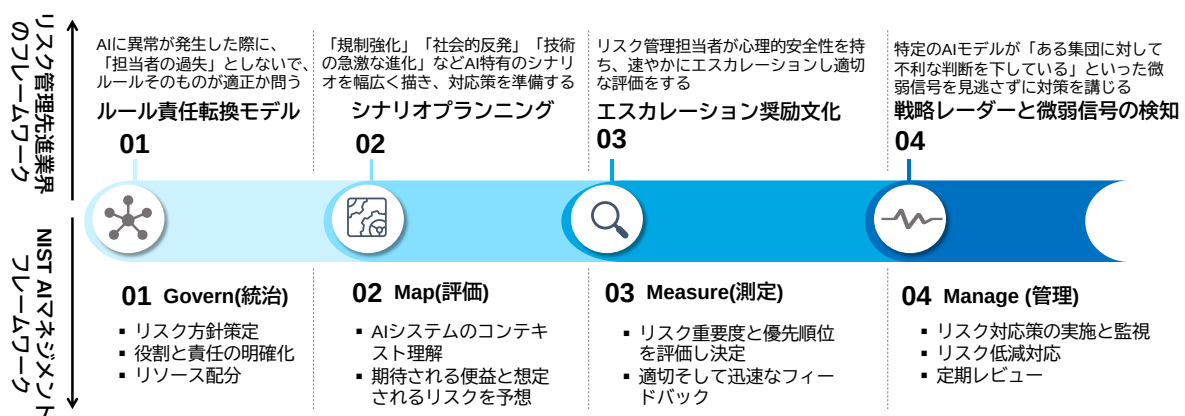
(2) AI リスク管理におけるフレームワークの提案

以上の管理手法を、時系列で企業活動における AI リスク管理に適用する：

1. 個人ではなく、リーダーが率先してルールを改善する仕組みを作り（ルール責任転換モデル）
2. 将来の多様なシナリオを想定して対応策を事前に練り（シナリオプランニング）
3. 組織内でリスクを上位に速やかに共有し（エスカレーション奨励文化）
4. 外部環境をリアルタイムに監視して前兆を検知し対応する（戦略リーダーと微弱信号の検知）

ことが推奨される。つまり、未来に起こりうるリスクをシナリオとして考えて、事前に対応策を備えることで、環境変化に対して極めて迅速な対応を起こすのだ(図表 8)。

図表 8：AI リスク管理フレームワーク



(出典) NIST(National Institute of Standards and Technology:米国国立標準技術研究所)

AI マネジメントフレームワーク^[20]をもとに筆者作成

以上の AI リスク管理フレームワークは、AI ビジネスの発展が爆発的スピードであることを踏まえると、極めて有効である。さらに、これは、NIST が定めた AI リスク管理フレームワーク^[20]に当てはめることができ、その信頼性は非常に高い。各業界が積み上げた未来への対応力を底上げするメソドロジーを適用することで、発生する危機に対して柔軟に対応することができるであろう。

6. 最後に：未知の脅威に備えるために他領域の知恵を活かす

AI は社会や産業構造を大きく変革する可能性を秘めた技術である一方、そのリスクやインパクトを正確に測ることは容易ではない。歴史が示すように、人類は大惨事の兆候を認知していても、楽観バイアスや正常性バイアスなどの心理的要因によって備えを怠りがちである。しかし、**未知のリスクに対する準備は、エネルギー、軍事、金融といった他領域でも行われてきた。**ここから得られた知見や知恵を、早急に AI リスク管理へ活用することが不可欠だ。

今度こそ、人類は過去の失敗や他業界の成功事例から学び、未知の大惨事を未然に防ぐ責務がある。AI が引き起こすかもしれない未曾有の危機を回避するためにも、私たちは積極的に他領域のリスク管理の知恵を取り入れ、強固な AI ガバナンスを構築していく必要がある。

参考文献

- [1] The White House. (2025, January 23). REMOVING BARRIERS TO AMERICAN LEADERSHIP IN ARTIFICIAL INTELLIGENCE, EXECUTIVE ORDER. Retrieved from The White House: <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/>
- [2] Kelly Ng, Brandon Dorennon, Tom Gerken, Marc Cieslak. (2025, February 5). DeepSeek: The Chinese AI app that has the world talking. Retrieved from BBC: <https://www.bbc.com/news/articles/c5yv5976z9po>
- [3] Kevin Okemwa. (2025, February 5). OpenAI's new "Deep Research" blows ChatGPT o3-mini and DeepSeek out of the water with 26.6% accuracy in the world's hardest "AI exam" — but it skipped the line. Retrieved from Windows Central: <https://www.windowscentral.com/software-apps/openai-deep-research-blows-chatgpt-o3-mini-and-deepseek-out-of-the-water>
- [4] SoftBank. (2025, February 3). OpenAI およびソフトバンクグループが提携し、企業用最先端 AI を開発・販売することに合意. Retrieved from SoftBank: https://www.softbank.jp/corp/news/press/sbkk/2025/20250203_01/
- [5] 山口真一. (2025, February 7). AI 監視枠組み、G7 が運用開始 開発企業に安全性報告促す. 日本経済新聞.
- [6] Wall Street Journal. (2025, February 12). JD Vance Tells World Leaders to Keep AI Regulation Light. Retrieved from: The Wall Street Journal: <https://www.wsj.com/video/jd-vance-tells-world-leaders-to-keep-ai-regulation-light /20F98802-327D-4F1C-90D3-E097925D1B39?msocid=3a6a10f3782164d808290318797c6596>
- [7] 経済産業省. (2024, November 22). AI 事業者ガイドライン(第 1.01 版). Retrieved from: 経済産業省. https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/20240419_report.html
- [8] Cornelius, P., Putte, A. V., & Romani, M. (2005, Fall). Three Decades of Scenario Planning in Shell. California Management Review, 48(1), pp.92-109.
- [9] Schoemaker, P. J., Day, G. S., & Snyder, S. A. (2013). Integrating organizational networks, weak signals, strategic radars and scenario planning. Technological Forecasting & Social Change, 80(4), pp.815-824.
- [10] Oil & Gas IQ Editor. (2019, October 19). Human Performance and Risk Management at BP. Retrieved February 13, 2025, from Oil & Gas iQ: <https://www.oilandgasiq.com/operational-excellence/interviews/human-performance-and-risk-management-at-bp>
- [11] Buehler, K., Freeman, A., & Hulme, R. (2008). The Risk Revolution. New York: McKinsey & Company.
- [12] UKEssays.Com. (2017, April 12). Audit and Analysis of Goldman Sachs. Retrieved from UKEssays.Com. <https://ukessays.com/essays/business/audit-and-analysis-of-goldman-sachs-business-essay.php>

- [13] Abigail Hofman. (2009, May 1). Stay with the strong. Retrieved from EUROMONEY:
<https://www.euromoney.com/article/27bjsstsqxhk1s1a0pb/opinion/abigail-hofman-stay-with-the-strong>
- [14] Wikipedia. (n.a.). Citibank. Retrieved from: Wikipedia:
<https://en.wikipedia.org/wiki/Citibank>
- [15] UBS. (2008). Fourth Quarter 2008 Results. Retrieved from UBS:
<https://www.ubs.com/global/en/investor-relations/financial-information/quarterly-reporting/qtr-archive/qtr-archive08-10.html>
- [16] RNZ. (2009, February 6). Loss reported by Deutsche Bank. Retrieved from RNZ:
<https://www.rnz.co.nz/news/business/40603/loss-reported-by-deutsche-bank>
- [17] Kevin Buehler, Andrew Freeman and Ron Hulme. (2008, September 1). The New Arsenal of Risk Management. Harvard Business Review.
- [18] Jorion Philippe. (2009). Risk Management Lessons from the Credit Crisis. European Financial Management, pp.1-19.
- [19] University of Cambridge. (2019). Cambridge Taxonomy of Business Risks. Cambridge: Cambridge Centre for Risk Studies.
- [20] NIST: National Institute of Standards and Technology. (2024). AI Risk Management Framework. Maryland: NIST.



東京海上ディール株式会社

CDO ユニット チーフデジタルオフィサー 大野 有生（専門分野：AI・デジタル活用）
〒100-0004 東京都千代田区大手町 1-5-1 大手町ファーストスクエア ウェストタワー 23F
Tel. 03-5288-6580 Fax. 03-5288-6590 | <https://www.tokio-dr.jp/>