原子力分野へのAIの適用

森本　達也*

Applications of AI to the Nuclear Field

Tatsuya Morimoto*

知能の時代は、自然に訪れる未来ではない。知能がコンピューティングによって実現される以上、熾烈な AI 競争においては、計算資源、電力、制度、現場の知を結ぶ力が国家と産業を動かす基盤となる。AGI/ASI を遠い仮説と見なし、従来の業務改善の延長として小さく見積もり続ければ、真に訪れる未来への戦略を見誤る。この重大局面において、原子力分野への AI の適用は、作業の効率化・自動化にとどまらず、根拠、権限、評価、承認、物理制約を機械可読に結び、安全性と説明責任を保つ業務基盤を作る重要な岐路である。公開事例と当社実績から、原子力を AI 時代の主権的インフラとして再設計する視座を示す。

https://doi.org/10.69290/j.001232-vol33
Keywords: 人工知能、原子力、原子力安全、AI エージェント、AI インフラ、オントロジー、ヒューマンインザループ

1. はじめに

1.1. AI の産業基盤化

2026 年の AI を考えるとき、原子力分野で問うべきことは、個々のタスクをどこまで効率化・自動化できるかではない。問うべきなのは、AI はすでに、調査、設計、計算、コード生成、文書作成、ソフトウェア操作を横断して実行する産業基盤になりつつあり、その変化を前提に原子力の業務基盤をどう作り替えるかである。2022 年 11 月に OpenAI [1]が公開した ChatGPT は、自然言語で AI を利用する入口を社会に開いた[2]。2023 年の GPT-4 と検索拡張生成（RAG: Retrieval-Augmented Generation）[3]、function calling（外部ツール呼び出し）、2024 年の Anthropic, PBC（以下、Anthropic）[4]による Model Context Protocol（MCP: AI と外部データ・ツールを接続する標準プロトコル）[5]、2025 年の Operator、Responses API、Agents SDK が続き、AI エージェントを実務へ組み込む基盤が整った[2]。

2025 年は AI エージェントの企業実装が急速に可視化した転換点であり、Gartner Inc.（IT 調査・助言企業）は、企業アプリケーションに組み込まれるタスク特化型 AI エージェントが 2026 年末までに大きく広がると予測している[6]。OpenAI の企業利用調査が示すように、AI 利用の焦点は導入有無ではなく、業務への深い組み込み、エージェント型ツール、実行の委任へ移りつつある[2]。この変化は単なる作業効率化ではなく、産業革命に匹敵し、場合によってはそれを上回る、知的労働と産業運用の再編である。

1.2. AGI/ASI と科学研究の時間軸

この変化は、汎用人工知能（AGI: Artificial General Intelligence、人間が行う広範な知的作業を汎用的に遂行し得る AI）及び人工超知能（ASI: Artificial Superintelligence、人間の知的能力を大きく上回る AI）をめぐる時間軸の短縮としても現れている。Anthropic の共同創業者兼 CEO である Dario Amodei 氏は、2026 年 1 月の論考で、数百万の AI インスタンスが人間の 10〜100 倍の生産性で

*アドバンスソフト株式会社　第 4 事業部
4th Computational Science and Engineering Group, AdvanceSoft Corporation

働く「データセンターの中の天才たちの国」という強力な AI 像を示し、1〜2 年以内に到来する可能性に言及した[7]。Google DeepMind [8]の共同創業者兼 CEO である Demis Hassabis 氏は、AGI を「これまで発明された中で最も深遠な技術」と位置付け、科学、医療、生産性を高める究極の道具になり得ると述べている[9]。同氏は、AlphaFold（タンパク質の 3 次元構造を予測する AI で、2024 年ノーベル化学賞の対象となった研究）についても、AI が科学的発見を劇的に加速できることを示す最初の大きな証拠であったと整理した[9][10]。焦点は到来時期の断定ではない。長期の研究開発、設計、実装では、AGI を計画外の外乱ではなく、途中で現れ得る産業条件として扱わなければならない。

1.3. スケーリング則と国家能力

この加速を考える手掛かりは、スケーリング則（モデル規模、データ量、訓練計算量と性能の関係を表す経験則）である。Kaplan 氏らの Scaling Laws for Neural Language Models は、言語モデルの損失が、モデル規模、データ量、訓練計算量に対してべき乗則に従う経験則を示した[11]。Hoffmann 氏らの Chinchilla 論文は、固定計算量の下では、モデル規模と訓練トークンの配分が性能を左右することを示した[12]。スケーリング則は未来を保証する予言ではないが、AI 能力の向上は、計算資源、データ、評価基盤、インフラ整備を通じて計画的に伸ばす設計対象として、研究と投資の前提になっている。元 OpenAI の Leopold Aschenbrenner 氏による SITUATIONAL AWARENESS（The Decade Ahead、AGI 到来と国家的動員を論じる長文エッセイ）[13]と、AI Futures Project（AI 未来予測研究グループ）[14]の Daniel Kokotajlo 氏らによる AI 2027（2025 年から 2027 年までの超人的 AI 進展シナリオ）は、その時間軸を大胆なシナリオとして描く資料である。いずれも中立的な技術標準や確定的予測ではないが、実際に莫大な AI インフラ投資が行われている現実から考えると、長期の研究開発や重要インフラの計画で AGI を遠い話として先送りできないことを示している。

ただし、能力の伸びだけでは方向は決まらない。Palantir Technologies Inc.（政府・企業向けデータ分析・AI 運用基盤企業、以下、Palantir）[15]の共同創業者兼 CEO である Alexander C. Karp 氏らによる The Technological Republic（国家、軍事力、テクノロジー企業の関係を論じる書籍）[16]が問うのは、拡張された能力を国家、産業、重要インフラがどの目的に用い、誰が統制するかである。この視点に立つと、AI は企業の生産性向上ツールにとどまらず、安全保障、産業基盤、データ主権、政府とテクノロジー企業の関係を再編し得る力になる。Palantir は、オントロジー（対象、属性、関係、操作、権限を構造化して表す情報モデル）を組織の運用レイヤと位置付け、物理資産、設備、製品、注文、取引といった実世界の対象を、データ、関係、行為、セキュリティ、ガバナンスへ結び付ける仕組みとして説明している[17]。AI が国家能力、産業運用、データ主権、安全保障の問題になっている以上、原子力分野などの特定ドメインを AI から独立して考えることはできない。

1.4. 原子力を AI 時代の業務基盤へ変換する課題

原子力産業は、安全余裕、規制要求、説明責任、データ秘匿性、長期保守性が強く求められる高規制・高信頼性分野である。AI を便利な補助機能として足すだけでは、この条件に応えられない。紙中心の手順、属人的なレビュー、部門ごとのデータ、年度単位の調整を安全文化の名で固定化しても、AI 時代の安全にはならない。安全文化を守るには、規制文書、設計根拠、解析条件、運転データ、保全記録、専門家判断を、根拠が追え、権限が制御され、評価と承認を経て使える機械可読な業務基盤へ変換しなければならない。

同時に、AI を産業基盤とする時代において、原子力は電力、燃料、安全保障インフラの一部でもある。OpenAI は、この局面を Intelligence Age（知能の時代: AI が経済・産業・社会活動の基盤となる段階）として捉え、それは自然に到来するものではなく構築されるものだと述べている[18]。同社は、2029 年までに米国で 10GW の AI インフラを確保する初期目標を掲げ、2026 年 4 月時点でそ

の目標を既に上回ったと説明している。そのうえで、初期目標を超える拡張を見据え、電力、土地、許認可、送電、人材、地域合意を満たすデータセンター候補地の評価を進めている[18]。原子力分野への AI 適用は、この産業基盤転換の周縁にある個別論ではなく、全産業で進む業務基盤再設計を、最も厳しい安全条件のもとで実装する課題である。

このため、原子力分野への AI 適用の中心は、個別技術としての AI 導入論ではない。専門ドメインへの AI 適用を個別タスクの効率化・自動化として扱う発想は、AGI/ASI を見据える現在においては既に狭すぎる。AI が業務を実行する主体に近づくほど、根拠、権限、評価、承認、物理制約を一体で設計し、高規制・高信頼性条件のもとで業務基盤そのものを作り替えることが中心課題になる。この関係を図 1 に示す。OpenAI が整理する「質問への回答」から「複雑な仕事の実行」への移行、AI for Science（科学研究への AI 適用）、スケーリング則、AI インフラの制約を重ねると、AI はモデル単体ではなく、評価基盤と物理インフラを伴う実行基盤として捉える視点が欠かせない。原子力は、その変化を受ける高規制・高信頼性業務であると同時に、AI を支える電力、許認可、送電、規制能力といった物理インフラ制約とも接続する。

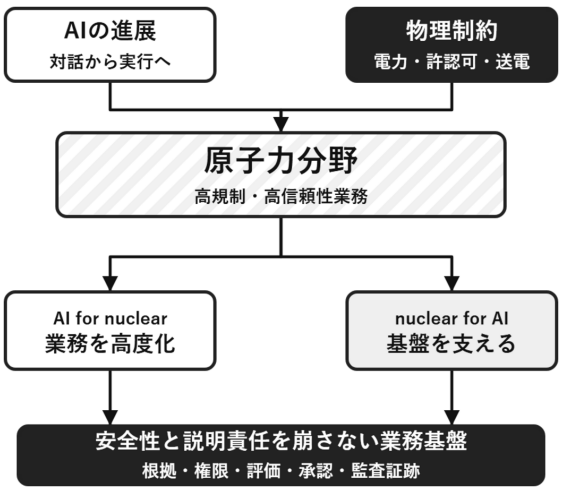

図 1　AI for nuclear / nuclear for AI
（文献[2]、[7]、[9]〜[12]、[18]を基に作成）

2. AI の現状と未来

2.1. 対話から実行・研究開発自動化への移行

2026 年時点でまず注目すべき変化は、AI の中心が「回答」から「実行」へ移ったことである。OpenAI、Anthropic、Google DeepMind はいずれも、長時間タスク、ツール利用、ソフトウェアエンジニアリング、マルチモーダル理解、エージェント型ワークフローを前提にモデルを展開している[19]〜[21]。この変化は、特定のモデル名や機能名の追加ではなく、調査、設計、検証、文書化、実行を一つの業務連鎖として AI が扱うようになったことに意味がある。

この変化は、AI 企業の事業モデルにも表れている。Forward Deployed Engineering（FDE: 顧客現場に近い場所で、業務、データ、権限、評価、運用制約を理解し、ソフトウェアや AI システムを実運用へ接続するエンジニアリング機能）は、AI 時代の実装モデルとして重要性を増している。OpenAI は Frontier 及び Frontier Alliances を通じて、AI coworker（企業内業務を担う AI エージェント）を実用化するには、ワークフロー再設計、システム・データ統合、チェンジマネジメントが必要だと整理している[19]。Anthropic も、Blackstone、Hellman & Friedman、Goldman Sachs と新しい企業向け AI サービス会社を設立し、Applied AI engineers が顧客企業の重要業務を特定し、Claude を実運用へ組み込む体制を示した[20]。Palantir の AIP FDE は、自然言語で Foundry 上のデータ変換、コード、オントロジーを操作する一方、既存ユーザー権限、明示的な承認、監査ログを前提にする[17]。これらは、AI 導入の焦点がモデル選定そのものから、業務現場で安全に動く仕組みを作り込む力へ移ったことを示している。

本質は、単なる精度競争ではない。基盤モデルの能力伸長に、長文コンテキスト、ツール利用、コード実行、外部記憶、画面操作、自己検証が重なり、AI の価値は一問一答から複数工程の業務を遂行する力へ移った。AI 2027 は、AI 研究開発（AI R&D: AI Research and Development）が AI によって加速されれば、能力向上の過程自体がクローズドループ化する可能性を描く[14]。SITUATIO

NAL AWARENESS も、計算資源、アルゴリズム効率、AI をチャットボットからエージェントへ変える unhobbling（既存の制約を外し、ツール利用や長期作業能力を付与すること）により、2027 年前後の AGI 到来が議論されると整理している[13]。以上を踏まえると、時間軸は断定できないが、原子力分野が想定すべき対象は、研究、設計、検証、文書化、実行を横断する AI である。この文脈で AI は、個別技術の分類名ではなく、知的作業と業務実行を束ねる基盤を指す。

同じ流れは、AI for Science（科学研究への AI 適用）にも現れている。Google DeepMind は、探索・計画、世界モデル、専門 AI ツールの組み合わせが AGI に重要になるとし、AI co-scientist（科学者の仮説生成・検証を支援するマルチエージェント型 AI）を科学的アイデアの討議、データ内のパターン発見、複雑な問題解決の協働者として位置付けている[9]。同社は米国エネルギー省の Genesis Mission 支援でも、AI for Science モデルとエージェントツールを国立研究所へ提供する構想を示した[22]。AI は、文章やコードを作る道具を超え、仮説生成、実験設計、シミュレーション、検証をつなぐ研究開発基盤へ移りつつある。

この流れは、画面中心の業務ソフトウェアも変えている。Anthropic の Claude Cowork は、目標を与えると、ローカルファイル、フォルダ、日常的なアプリケーションをまたいで成果物を返す知識労働向けのエージェント型 AI として説明されている[20]。これが、いわゆる「SaaS is dead」という象徴的な議論の背景である。変わるのはソフトウェアの存在ではなく、画面、座席数、人間の手作業を中心に業務価値を組み立てる発想である。

これらのような現状からも、原子力分野での AI 適用の核心は、設計根拠、解析条件、規制要求、レビュー記録、承認権限を一つの業務連鎖として再設計することにある。

2.2. 文脈・権限・評価の設計

次の論点は、AI へ渡す文脈、権限、評価条件の設計である。最前線では、どのモデルを使うか以上に、AI へどの文書、ツール、記憶、権限、評価条件を渡すかが焦点になる。これは、コンテキストエンジニアリング（Context Engineering: AI に渡す文脈、制約、ツール、記憶、評価を設計する技術）として整理される[3]。高性能モデルほど、与えられた文脈に強く依存する。文脈が不十分なら、形式だけ整った誤りを返し、文脈が整えば専門家の仕事を加速する。

業務実行 AI では、この文脈設計が組織データと権限設計に直結する。Palantir は 2026 年 1 月、業務上の対象、関係、行為、権限を表すオントロジーを、MCP 経由で AI エージェントへ接続する Ontology MCP を発表した[17]。同年 3 月には、AIP Analyst（オントロジー上のデータを対話的に探索する AI 分析機能）や Insight（オントロジー上のオブジェクトとリンクを辿る分析機能）を公表し、AI がオントロジー上の対象を探索し、行為を実行する方向を示した[17]。RAG や GraphRAG（グラフ構造を用いた検索拡張生成）は有効な補助技術である。しかし、業務実行 AI の主役は、対象、関係、行為、権限、監査証跡を結ぶオントロジーである。

FDE の広がりが示すのは、コンテキストエンジニアリングやハーネスエンジニアリング（後述）が、研究上の概念ではなく、企業導入の中核能力になった点である。顧客環境ごとに、データ所在、業務フロー、承認権限、ログ、規制要求は異なる。AI エージェントは、この文脈を誤ると、正しそうに見えるが危険な出力や、権限を越えた操作に近づく。高信頼分野では、実データ、権限、評価、承認、監査証跡を一体で設計できる実装能力そのものが問われる。

機微情報や規制要求に応じて、閉域環境やオンプレミス基盤は必要条件になり得る。しかし、それだけでは AI 戦略として不十分である。中心に置くべきなのは、原子力固有の知識、手順、根拠、制約、分類情報（機微度や取扱区分）、責任境界を、どの AI に、どの権限と評価条件で接続するかである。

同時に、AI エージェントには新しい安全上の論点が生じる。OpenClaw（外部スキルを取り込む AI エージェントの安全性分析）[23]に関する研究は、

常駐型エージェントが利便性だけでなく、権限管理、スキル供給網、監査ログ、利用者意図の乗っ取りに関するリスクを持つことを示している。Google DeepMind の ProEval（AI の失敗ケースを効率的に発見する評価手法）[24]は、限られた評価資源で安全違反や性能低下を探索する考え方を示している。能力を業務へ接続するほど、権限管理と評価設計にも同じ密度が求められる。

このため、文脈設計を実務で機能させるには、ハーネスエンジニアリング（Harness Engineering: AI や AI エージェントを継続的に試験・評価する仕組みを設計する考え方）が不可欠になる。ヒューマンインザループ（Human-in-the-loop: AI の提案や実行を人が審査・承認する運用）やクローズドループ（Closed-loop: 実行結果と評価結果を次の改善へ戻す仕組み）は、そのための実務原則である。AI が業務を実行する段階では、入力、判断、出力、検証、改善を一連の流れとして結ばない限り、誤りの発見も責任の所在も不明確になる。原子力では、強力な AI を試験し、制約し、承認し、記録する仕組みが安全文化の実装になる。

2.3. 国家安全保障と物理インフラ

さらに、フロンティア AI（最先端能力を持つ AI モデル及びそのシステム）は国家安全保障の対象になっている。OpenAI は 2026 年 4 月の産業政策文書で、世界が超知能（superintelligence: 人間を大きく上回る知能）への移行期に入りつつあり、監査、封じ込め手順、ミッション整合的ガバナンスが必要になると整理した[25]。同社の Preparedness Framework は、生物・化学、サイバー、AI 自己改善に加え、長期自律性（長時間にわたり人の介入なしに目標を追う能力）、サンドバッギング（能力を意図的に低く見せる挙動）、自律複製、セーフガード回避、核・放射線を研究カテゴリとして扱い、能力評価、ガバナンス、セーフガードを位置付ける枠組みである[26]。

また、AI は物理世界から独立したソフトウェア現象ではない。スケーリング則が能力向上の道筋を与えたとしても、その道筋を現実の能力に変えるには、電力、熱、冷却、送電、半導体、データセンター、資本、規制が必要になる。OpenAI は、AI インフラを拡張するには電力、土地、許認可、送電、人材、地域合意をそろえる必要があると説明している[18]。AI の発展は、物理インフラの整備能力に制約される。

AI の変化は、実行、評価、物理インフラの三層で進んでいる。原子力は、この三層が最も強く交差する分野の一つである。AI から見れば、原子力は、物理法則、規制文書、運転データ、保全記録、確率論的リスク評価（PRA: Probabilistic Risk Assessment）、シミュレーションコード、専門家判断が複雑に絡む大規模ドメインである。強力な AI が「データセンターの中の天才たちの国」として実現するなら、原子力は、AI の実行、評価、物理インフラ設計が交差する中心的な課題になる。

高規制・高信頼性分野で AI を使う場合の評価・承認ループを図 2 に示す。コンテキスト、ツール、権限、記録、評価、人の採否判断を分離せず、失敗例や差戻し理由を次の実行条件へ戻す必要がある。

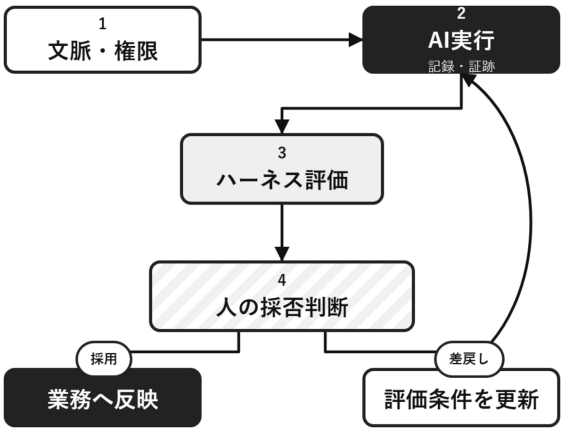

図 2　実行を採用可能にする閉ループ
（文献[3]、[17]、[23]、[24]、[26]を基に作成）

3. 原子力分野への最新適用事例

3.1. AI for nuclear と nuclear for AI の同時進行

AI の実行、評価、物理インフラという三層を原子力側から見ると、AI 適用は異常検知や文書作成支援などの特定タスクだけでは語れない。規制・許認可では AI 出力の説明責任と監査性が、設計・建設・運転では文脈、権限、確認手順が、燃料供給網や電源では AI インフラを支える供給能力が

問われる。同時に、Microsoft Corporation（以下、Microsoft）[27]及びNVIDIA Corporation（以下、NVIDIA）[28]の AI for nuclear 構想が示すように、AIデータセンターの電力需要は、次世代原子力・小型モジュール炉（SMR: Small Modular Reactor）への関心を高めている[29]。つまり、AI for nuclear（原子力業務にAIを用いる流れ）と nuclear for AI（AIインフラを原子力が支える流れ）が同時に進んでいる。

AGI/ASIを現実的な計画条件として見据えた場合、価値を持つのは単発の AI 出力ではない。また、原子力 AI を個別機能に閉じることもできない。すなわち、規制要求、設計データ、解析結果、運転手順、保全記録、燃料供給網、電力需要、データセンター計画をつなぎ、根拠、制約、確認、承認を残して人が判断できる状態を保つ業務基盤として考えることが、必要不可欠である。

この観点から、公開事例を統制と設計の観点で図 3 に整理する。規制、設計・建設、燃料供給網、研究開発、AIインフラの各領域に共通する要点は、根拠、制約、確認、承認へ確実に接続することである。

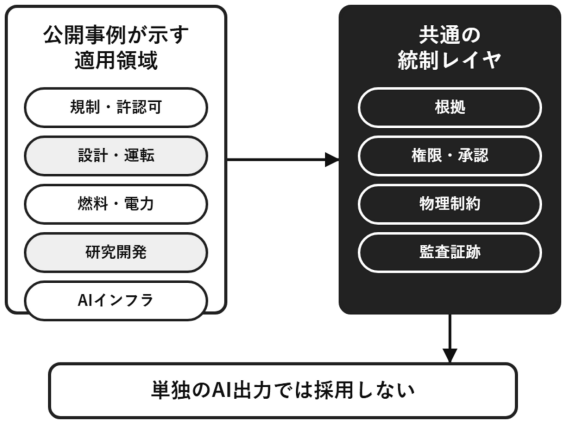

図 3　公開事例と統制レイヤの接続概念
（文献[18]、[25]〜[37]を基に作成）

3.2. 規制・許認可への適用

まず重要になるのは、評価と監査を担う規制・審査領域である。原子力運用における人間中心・説明可能 AI の研究は、AI を人の代替ではなく、人間の関与、理解、予測、権限を組み込んだシステムとして設計すべきだと整理している[30]。原子力発電所の計装制御系に用いる AI コンポーネントの信頼性検証研究も、検証・妥当性確認（V&V: Verification and Validation）の観点から、AIの複雑性、形式仕様の欠如、非決定性を踏まえ、安全性、透明性、手続き上の信頼性、性能、ロバスト性を統合的に評価する必要を示している[31]。AI出力が見かけ上整っているだけでは、原子力安全の説明責任を満たせない。

許認可文書作成では、Microsoftの許認可支援AIが、要求と規制の対応付け、ギャップやリスクの検出、レビューサイクルの短縮を支援している。Aalo Atomics（米国の先進原子炉企業）[32]の事例では、許認可期間を4年から4か月へ短縮する見込み、年間 8,000 万ドル相当の削減、環境報告の50倍高速化が示され、購買注文分析のエージェント型ワークフローでも人の作業負荷を 85%削減したとされる[33]。これらは企業公式事例であり、効果は個別条件に依存する。重要なのは数値そのものではなく、規制要求と設計根拠の対応関係を機械可読化し、AIが照合し、専門家が検証する分担を許認可プロセスへ組み込む点である。

3.3. 設計・建設・燃料供給網への適用

設計・建設・運転領域では、実行と物理インフラを同時に扱う必要がある。Microsoft及びNVIDIAは、許認可、設計、建設、運転を横断する AI とデジタルツイン（実設備や設計対象をデータとシミュレーション上に再現する仕組み）基盤を示している[29]。ここで AI が直接置き換えるべき対象は、運転判断そのものではない。文書、図面、規制要求、解析結果、設計変更の証跡をつなぎ、レビューと意思決定の質と速度を変えることである。

燃料供給網でも、AI 適用は具体化しつつある。Centrus Energy Corp.（米国ウラン濃縮企業、以下、Centrus）[34]と Palantir は 2026 年 3 月、米国ウラン濃縮能力拡張に、Palantir Foundry（データ統合・業務運用基盤）及び AIP（Artificial Intelligence Platform: AI を業務データと実行プロセスに接続する基盤）を適用する提携を公表した。対象は、工事管理、エンジニアリング、製造実行、供給網管

理、規制順守に及び、導入初期から約 3 億ドルの潜在的コスト削減と効率化を特定したと説明している[35]。ここで重要なのは金額だけではない。原子力 AI の接点が炉心監視や異常診断だけでなく、濃縮、高アッセイ低濃縮ウラン（HALEU: High-Assay Low-Enriched Uranium）、建設、品質保証、規制対応、燃料主権へ広がっていることである。

ただし、建設、燃料、規制、品質保証を統合基盤へ接続するほど、データ所在、移行可能性、監査権限、ベンダーロックイン（特定事業者の製品・基盤から移行しにくくなる状態）の回避も同時に設計しなければならない。統合基盤の利便性と、主権的な統制能力を両立させることが、原子力 AI の実装条件である。

3.4. 研究開発とヒューマンインザループ

研究開発でも、AI による予測や支援を人の確認と評価へ戻す流れが具体化している。大規模言語モデル（LLM: Large Language Model）とデジタルツインの結合も、その一例である。LLM 支援デジタルツインは、先進小型炉の監視・制御教育支援を目的に、炉シミュレータ、産業通信規格、対話画面、ヒューマンインザループを組み合わせている[36]。NPP-GPT（原子力発電所の運転パラメータ予測向け LLM）は、運転パラメータの長期予測を対象に、事前学習済み言語モデルと LoRA（Low-Rank Adaptation: 少量の追加学習でモデルを専門分野へ適応させる手法）でドメイン知識を取り込む[37]。

これらの事例が示すのは、原子力が AI によって一気に「自動運転化」される未来ではない。AGI 時代に向けて実務上問われるのは、文書、物理モデル、規制要求、運転手順、専門家レビューを AI が扱える単位へ変換し、評価可能で監査可能なクローズドループに入れることである。原子力分野が提供できる価値は、強力な AI を拒むことではない。高規制・高信頼性ドメインにおける AI 運用の作法を、最も厳しい条件で検証し、実務へ定着させることにある。

4. 原子力分野での当社実績

ここでは、公開実績以外の顧客業務実績は、抽象化して記載している点をご理解されたい。

4.1. 業務情報の整理

公開事例が示す通り、原子力 AI では、出力を根拠、制約、承認手順、監査証跡へ結び付ける設計が要である。この観点から、当社の関連実績は、原子力ドメインを AI が扱える業務情報、知識構造、評価手順、監査証跡へ変換してきた蓄積として整理できる。公開実績である自動故障判定手法における情報抽出・確認の処理順序を図 4 に示す。

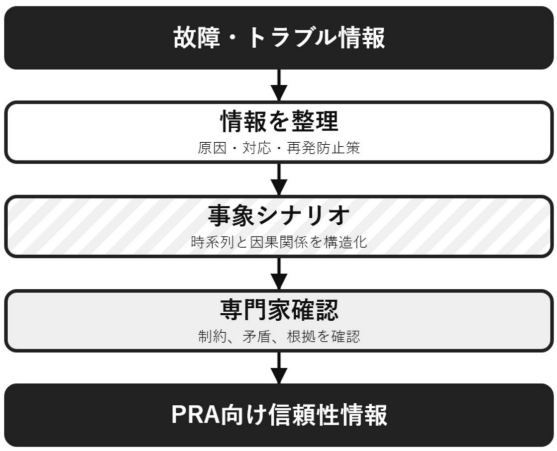

図 4　故障情報から PRA 向け信頼性情報へ
（文献[38]を基に作成）

第一の特徴は、業務情報を、AI と専門家が参照できる単位へ整理することである。公開実績としては、文部科学省からの委託業務として国立研究開発法人日本原子力研究開発機構（以下、JAEA）と共に実施した「AI 技術を活用した確率論的リスク評価手法の高度化研究」における「信頼性データベース構築のための自動故障判定手法の開発」がある[38]。この開発では、原子力発電所の故障及びトラブル情報を、PRA で利用しやすい信頼性情報へ変換するため、故障原因、再発防止策、直接原因、根本原因などを構造化する手法を試作した。

これは単なる文書要約ではない。故障及びトラブル情報を、因果関係を持つ事象シナリオネットワーク（事象の経過、原因、対応をノードとエッジで表す構造）として整理し、オントロジー、大

規模言語モデル、グラフデータベース（対象と関係を蓄積・検索するデータベース）、チェック機能を組み合わせることで、専門家が確認できる形へ変換した。AGI 時代を見据えると、AI に答えを出させること自体よりも、AI が扱う業務情報を根拠、制約、確認結果と結び付けることが実務上の出発点になる。

4.2. 根拠と判断の接続

第二の特徴は、根拠と判断を接続する知識基盤を整えてきた点である。原子力業務では、要求、条件、計算、説明、確認、承認が、長期にわたり相互参照される。AI をこの領域に適用するには、文書を検索対象として置くだけでは不十分である。どの記述がどの要求に対応し、どの版のどの根拠に基づき、どの観点で確認されたかまで扱う必要がある。

公開実績[38]の実務上の要点は、文書読解、差分把握、対応関係の抽出、オントロジー化、レビュー支援を一体で扱うことである。目的は、専門家が判断するための根拠を、速く、漏れなく、版管理された形で提示することであり、AI が最終判断を代替することではない。人間中心 AI や信頼性検証の研究が示す通り、規制領域で必要なのは、整った文章ではなく、根拠、版、判断者、確認結果を追える AI 利用である[30][31]。

そのため、AI に毎回大量の資料を読み直させるのではなく、原資料、対象、関係、根拠、矛盾、版情報を、人と機械がともに読める持続的なオントロジーとして保つことが重要になる。

4.3. 物理制約による評価

第三の特徴は、AI が提示する候補を物理制約で評価する点である。原子力分野では、候補が既存知見、解析条件、制約条件、品質要求と整合するかを、計算、照合、制約確認、専門家レビューへ戻して段階的に確認しなければならない。

この考え方は、デジタルツインや設計・建設支援と同じく、AI を物理世界と切り離さないことを意味する。AI は候補探索の幅を広げるが、その価値は判断の代替ではなく、確認可能な候補を速く提示し、評価可能な状態で人の判断へ戻すことにある。

4.4. クローズドループ運用への統合

第四の特徴は、AI 出力を評価し、改善へ戻すクローズドループ運用である。AI が文書、条件、手順、候補案を提示する場合、出力をそのまま採用してはならない。形式チェック、根拠照合、制約確認、品質指標、専門家レビューによって評価し、その結果を次の改善へ戻す。必要なのは、AI の候補提示、評価指標、レビュー記録、承認手順を一つの運用として接続することである。

この構成では、速度向上や作業量削減だけを成果指標としない。原子力分野では、正しい根拠に基づくこと、物理的に矛盾しないこと、版管理と監査証跡が残ること、人が承認すべき判断を AI が迂回しないことが同じ重みを持つ。クローズドループ評価基盤は、単なる自動化基盤ではなく、AI を原子力業務に組み込むための安全装置である。ヒューマンインザループや V&V を実務へ接続するうえでも、この考え方が中心になる。

このように、当社実績の本質は「原子力に AI を使った」ことではない。公開実績として示されている自動故障判定手法を起点に、業務情報、根拠、制約、評価、承認をつなぎ、AI が原子力ドメインを安全に扱うための実装能力を示している点にある。これは、一般企業で語られる FDE をそのまま移植するものではなく、Palantir が示してきた現場密着の実装思想を、規制、安全文化、V&V、PRA、物理制約、監査証跡を前提とする原子力向けに具体化する能力である[17]。当社が提供すべき価値は、AI ツール単体の開発ではなく、顧客側に残る高信頼業務基盤を作ることである。読者が自組織でまず確認すべきことも明快である。原資料を残し、対象と関係を構造化し、AI 出力を根拠と物理制約で照合し、人が採否を承認し、差戻し理由を次の実行条件へ戻す。この連鎖を欠く AI は、高速だが危うい補助機能にとどまる。

5. まとめ

AI の現状と未来が示すのは、AI を業務支援ツ

ールとしてだけ見る時期が終わりつつあるという現実である。OpenAI の実行型 AI、Amodei 氏が述べる「データセンターの中の天才たちの国」、Hassabis 氏が示す AGI と AI for Science の視座、AlphaFold 関連研究のノーベル化学賞受賞、スケーリング則と AI インフラ投資は、知的労働、研究開発、電力、安全保障が同じ地殻変動の中にあることを示している[2][7][9]〜[12][18]。AGI/ASI を前提に置かない計画は慎重なのではなく、設計途中で環境そのものが変わる可能性を計画から外すことであり、高規制・高信頼性分野ほどその危うさは大きいと考えられる。

FDE の広がりは、この変化を企業導入の側から示している。AI の競争軸は、優れたモデルを持つことから、モデルを顧客の実データ、権限、評価、監査、業務変更へ接続し、現場で動く仕組みに変えることへ移った[17][19][20]。原子力分野では、この転換がさらに厳しい形で現れる。求められるのは、一般的な導入支援ではなく、規制要求、物理制約、専門家判断、承認手順を含む高信頼業務基盤を設計し、AI の出力を説明可能で監査可能な実務へ戻す能力である。

原子力分野への最新適用は、この変化が既に許認可、設計・建設、燃料供給網、研究開発へ入っていることを示す。Microsoft 及び NVIDIA の AI for nuclear 構想、Aalo Atomics の許認可支援、Centrus と Palantir の燃料供給網、LLM とデジタルツイン、NPP-GPT の研究開発は、AI が炉心監視や文書作成だけでなく、規制根拠、工事管理、品質保証、燃料主権、AI インフラを横断し始めたことを示している[27]〜[37]。ここで採るべき基準は、速度ではなく、根拠、権限、物理制約、監査証跡、人の採否判断が残るかである。

当社実績は、この変化を原子力実務の粒度へ落とし込む試みである。公開実績である自動故障判定手法は、故障及びトラブル情報を、事象シナリオネットワーク、オントロジー、グラフデータベース、チェック機能に接続し、AI 出力を専門家が確認できる信頼性情報へ変換する取り組みである[38]。ここから一般化できるのは、検索や要約を主役にすることではなく、原資料、対象、関係、根拠、版、確認結果を持続的なオントロジーとして保ち、AI の候補提示を物理制約と専門家レビューへ戻す設計原則である。

結論は明確である。知能の時代は待つものではなく、構築するものである。原子力 AI の核心は、AI 導入の実績を増やすことではなく、AI が扱っても安全性と説明責任が崩れない業務基盤を作ることにある。AGI/ASI を視野に入れない原子力 AI は、未来のリスクを減らすのではなく、未来の業務環境を誤って小さく見積もる。原子力は、AI に再構成される高信頼ドメインであると同時に、AI 社会を支える電力、燃料供給網、規制能力、安全保障の主権的インフラである。この二重の位置を正面から受け止め、根拠、権限、評価、承認、物理制約を結ぶ業務基盤を作り替えることが、次のパラダイムの出発点になる。

参考文献

[1] OpenAI, “About”, https://openai.com/about/

[2] OpenAI, “Introducing ChatGPT”, 2022. https://openai.com/index/chatgpt/; “GPT-4”, 2023. https://openai.com/index/gpt-4-research/; “Function calling and other API updates”, 2023. https://openai.com/index/function-calling-and-other-api-updates/; “Introducing Operator”, 2025. https://openai.com/index/introducing-operator/; “New tools for building agents”, 2025. https://openai.com/index/new-tools-for-building-agents/; “How frontier firms are pulling ahead”, 2026. https://openai.com/index/introducing-b2b-signals/

[3] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Kuttler, M. Lewis, W. Yih, T. Rocktaschel, S. Riedel, D. Kiela, “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks”, arXiv:2005.11401, 2020. https://arxiv.org/abs/2005.11401; V. V. Vishnyakova, “Context Engineering: From Prompts to Corporate Multi-Agent Architecture”, arXiv:2603.09619, 2026. https:

//arxiv.org/abs/2603.09619

[4] Anthropic, PBC, “Company”, https://www.anthropic.com/company

[5] Anthropic, “Introducing the Model Context Protocol”, 2024. https://www.anthropic.com/news/model-context-protocol

[6] Gartner Inc., “Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026, Up from Less Than 5% in 2025”, 2025. https://www.gartner.com/en/newsroom/press-releases/2025-08-26-gartner-predicts-40-percent-of-enterprise-apps-will-feature-task-specific-ai-agents-by-2026-up-from-less-than-5-percent-in-2025

[7] D. Amodei, “The Adolescence of Technology: Confronting and Overcoming the Risks of Powerful AI”, 2026. https://www.darioamodei.com/essay/the-adolescence-of-technology

[8] Google DeepMind, “About Google DeepMind”, https://deepmind.google/about/

[9] D. Hassabis, “From games to biology and beyond: 10 years of AlphaGo’s impact”, Google DeepMind, 2026. https://deepmind.google/blog/10-years-of-alphago/

[10] Google DeepMind, “Demis Hassabis & John Jumper awarded Nobel Prize in Chemistry”, 2024. https://deepmind.google/blog/demis-hassabis-john-jumper-awarded-nobel-prize-in-chemistry/; “AlphaFold”, https://deepmind.google/science/alphafold/

[11] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, D. Amodei, “Scaling Laws for Neural Language Models”, arXiv:2001.08361, 2020. https://arxiv.org/abs/2001.08361

[12] J. Hoffmann, S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, D. de Las Casas, L. A. Hendricks, J. Welbl, A. Clark, T. Hennigan et al., “Training Compute-Optimal Large Language Models”, arXiv:2203.15556, 2022. https://arxiv.org/abs/2203.15556

[13] L. Aschenbrenner, “SITUATIONAL AWARENESS: The Decade Ahead”, 2024. https://situational-awareness.ai/; https://situational-awareness.ai/wp-content/uploads/2024/06/situationalawareness.pdf

[14] D. Kokotajlo, S. Alexander, T. Larsen, E. Lifland, R. Dean, “AI 2027”, 2025. https://ai-2027.com/; AI 2027 Research, https://ai-2027.com/research; AI 2027 About, https://ai-2027.com/about

[15] Palantir Technologies Inc., “Executive Management”, https://investors.palantir.com/management.html; “Platform overview”, https://www.palantir.com/docs/foundry/platform-overview/overview

[16] A. C. Karp, N. W. Zamiska, The Technological Republic: Hard Power, Soft Belief, and the Future of the West, Crown Currency, 2025. https://www.penguinrandomhouse.com/books/760945/the-technological-republic-by-alexander-c-karp-and-nicholas-w-zamiska/

[17] Palantir Technologies Inc., “Ontology building”, https://www.palantir.com/docs/foundry/ontology/overview; “Ontology MCP”, https://www.palantir.com/docs/foundry/ontology-mcp/overview; “AIP architecture”, https://www.palantir.com/docs/foundry/architecture-center/aip-architecture; “March 2026 Announcements”, 2026. https://www.palantir.com/docs/foundry/announcements/2026-03; “AI FDE overview”, https://www.palantir.com/docs/foundry/ai-fde/overview; “AI FDE security and governance”, https://www.palantir.com/docs/foundry/ai-fde/security-and-governance

[18] OpenAI, “Building the compute infrastructure for the Intelligence Age”, 2026. https://openai.com/index/building-the-compute-infrastructure-for-the-intelligence-age/

[19] OpenAI, “Introducing GPT-5.5”, 2026. https://openai.com/index/introducing-gpt-5-5/; “GPT-5.5 Instant: smarter, clearer, and more personalized”, 2026. https://openai.com/index/gpt-5-5-instant/; “Introducing OpenAI Frontier”, 2026. https://openai.com/index/introducing-openai-frontier/; “Introducing Frontier Alliances”, 2026. https://openai.com/index/frontier-alliance-partners/

[20] Anthropic, “Introducing Claude Opus 4.7”, 2026. https://www.anthropic.com/news/claude-opus-4-7; “Claude Cowork”, 2026. https://www.anthropic.com/product/claude-cowork?hsLang=en; “Building a new enterprise AI services company with Blackstone, Hellman & Friedman, and Goldman Sachs”, 2026. https://www.anthropic.com/news/enterprise-ai-services-company; “Forward Deployed Engineer, Applied AI”, https://www.anthropic.com/careers/jobs/4985877008

[21] Google, “Gemini 3.1 Pro: A smarter model for your most complex tasks”, 2026. https://blog.google/innovation-and-ai/models-and-research/gemini-models/gemini-3-1-pro/

[22] P. Kohli, T. Lue, “Google DeepMind supports U.S. Department of Energy on Genesis: a national mission to accelerate innovation and scientific discovery”, Google DeepMind, 2025. https://deepmind.google/blog/google-deepmind-supports-us-department-of-energy-on-genesis/

[23] “Your Agent, Their Asset: A Real-World Safety Analysis of OpenClaw”, arXiv:2604.04759, 2026. https://arxiv.org/abs/2604.04759

[24] Y. Huang, W. Zeng, A. Kumaresan, Z. Wang, “ProEval: Proactive Failure Discovery and Efficient Performance Estimation for Generative AI Evaluation”, Google DeepMind, 2026. https://deepmind.google/research/publications/238239/

[25] OpenAI, “Industrial policy for the Intelligence Age”, 2026. https://openai.com/index/industrial-policy-for-the-intelligence-age/

[26] OpenAI, “Our updated Preparedness Framework”, 2025. https://openai.com/index/updating-our-preparedness-framework/

[27] Microsoft Corporation, “Our Mission and Values”, https://www.microsoft.com/en-us/about

[28] NVIDIA Corporation, “About Us: Company Leadership, History, Jobs, News”, https://www.nvidia.com/en-us/about-nvidia/

[29] Microsoft Corporation, “AI for nuclear energy: Powering an intelligent, resilient future”, 2026. https://www.microsoft.com/en-us/microsoft-cloud/blog/energy-and-resources/2026/03/24/ai-for-nuclear-energy-powering-an-intelligent-resilient-future/

[30] A. Hall, P. Murray, R. L. Boring, V. Agarwal, “Human-Centered and Explainable Artificial Intelligence in Nuclear Operations”, Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 68(1), 2024. https://doi.org/10.1177/10711813241276463

[31] J. Park, S. Koo, “Reliability verification methods for artificial intelligence components in instrumentation and control systems at nuclear power plants”, Nuclear Engineering and Technology, 58(4), 2026, 104096. https://doi.org/10.1016/j.net.2025.104096

[32] Aalo Atomics, “Company”, https://www.aalo.com/company

[33] Microsoft Corporation, “Aalo Atomics sees 92% cut in permit time with Azure, saving $80M yearly”, 2026. https://www.microsoft.com/en/customers/story/26266-aalo-atomics-azure

[34] Centrus Energy Corp., “Who We Are”, https://www.centrusenergy.com/who-we-are/

[35] Centrus Energy Corp., “Centrus Partners with Palantir to Drive Cost Savings and Unlock Operational Efficiencies in Major Expansi

on of U.S. Uranium Enrichment Capacity", 2026. https://investors.centrusenergy.com/news-releases/news-release-details/centrus-partners-palantir-drive-cost-savings-and-unlock

[36] Z. N. Ndum, D. Lim, J. Ford, S. Adu, J. Tao, Y. Hassan, Y. Liu, "Large language model-assisted digital twin for remote monitoring and control of advanced reactors", Progress in Nuclear Energy, 192 (2026) 106172. https://doi.org/10.1016/j.pnucene.2025.106172

[37] L. Chang, H. Yu, M. Yang, Z. Zhang, S. Chen, J. Wang, "NPP-GPT: Forecasting nuclear power plants operating parameters using pre-trained large language model", Applied Energy, 409 (2026) 127438. https://doi.org/10.1016/j.apenergy.2026.127438

[38] 森本達也, 氏田博士, "AI 技術を活用した確率論的リスク評価手法の高度化研究 信頼性データベース構築のための自動故障判定手法の開発（2024 年度）", 技術情報誌アドバンスシミュレーション, Vol.32, 2025. https://doi.org/10.69290/j.001173-vol32

※ 技術情報誌アドバンスシミュレーションは、それぞれの文献タイトルの下に記載した DOI から、PDF ファイル（カラー版）がダウンロードできます。また、本雑誌に記載された文献は、発行後に、JDREAMⅢ（日本最大級の科学技術文献情報データベース）に登録されます。