

東大生研「MatAgent」と最新材料探索AI技術の詳細分析

要約 (Summary)

- **MatAgentの特徴:** 東京大学生産技術研究所の研究グループが開発した「MatAgent」は、大規模言語モデル (LLM) をエージェントとして用い、無機材料の組成を人間専門家のように推論しながら自律的に探索・設計するフレームワークです¹。LLMが短期記憶・長期記憶・周期表・知識ベースといった外部ツールを駆使し、各候補の提案理由を自然言語で説明することで、ブラックボックスになりがちな材料探索プロセスに高い**解釈性**を持たせています²³。
- **技術の詳細:** MatAgentでは汎用LLM (OpenAI GPT-4ベースのモデル等) を「**頭脳**」とし、**Planning** (計画) と**Proposition** (提案) の2段階で推論を行います。LLMは必要に応じて先述の外部ツールを呼び出し、過去の試行結果や材料知識を参照しながら新たな材料組成を提案します⁴。提案された組成は**拡散モデル**による3次元構造生成と、**GNNベース予測モデル** (形成エネルギー等の物性を評価) による性能評価を経て、フィードバックがLLMに返され次の反復に活かされます¹⁵。このループを通じ、LLMエージェントは徐々に目標特性に近づく材料を発見していきます。
- **関連研究動向:** 近年 (2024~2025年)、大手企業各社も生成AIやLLMを活用した材料探索に注力しています。Google DeepMindはグラフニューラルネットワークを駆使した**GNoME**により220万種を超える新規結晶材料 (うち38万種は安定構造) を発見しました⁶。Microsoftは**MatterGen**と呼ばれる拡散モデルで格子・原子配置・元素種を同時に生成し、AIが直接ターゲット物性に合致する構造を提案する手法を確立しています⁷⁸。IBMも化学領域で大規模**基盤モデル**を公開し、複数のデータ表現 (SMILES, グラフ等) を組み合わせたマルチモーダルAIで新材料探索の効率化を図っています⁹¹⁰。スタートアップ企業では、Kebotixのようにロボット実験とAIを組み合わせた自動化実験室や、Citrine Informaticsのように材料データベースと機械学習で特性予測・最適化を行うプラットフォームが登場しつつあります¹¹。
- **アプローチの比較:** MatAgentに代表される「**専門家のような推論型**」エージェントは、提案理由を示す**説明可能性**と、人間の暗黙知を取り入れた**探索プロセス**が特長です。一方、企業の「**大規模計算・データ駆動型**」アプローチは、大量の組成空間を網羅的に計算し高性能候補を抽出することに優れますが、結果の因果理解や柔軟な制約適用は必ずしも得意ではありません¹²。MatAgentは探索過程そのものから科学的知見を引き出しうる点で学術的意義が大きく、企業の効率重視手法と**競合**と言うよりは**補完的**な関係にあります。本報告では、これらの点を踏まえMatAgentの技術内容と企業各社の最新動向を詳細に分析し、産業界への示唆を考察します。

1. 東大生研・MatAgentの技術的特徴 (自律性・解釈性・専門家のような推論)

東京大学生産技術研究所の高原泉氏・溝口照康氏らのグループは、「**MatAgent**」と名付けたAIエージェント型フレームワークを開発しました¹³。MatAgentは**大規模言語モデル (LLM)** を中核に据え、専門家が材料を探すような推論プロセスをそのままAIに行わせる点が大きな特徴です。以下、MatAgentの重要な技術的特徴である「**自律性**」「**解釈性**」「**専門家のような推論プロセス**」について解説します。

- **自律性 (オートノマス探索):** MatAgentは、ユーザが設定した**目標特性**に向けて材料探索を**自律的 (オートノマス)**に進めます⁴。具体的には、LLMが試行錯誤を重ねながら新たな材料組成を次々

と提案し、各提案に対して外部モデルで構造生成・特性評価を行い、その結果（フィードバック）を受けて次の提案に活かすという**計画→提案→評価→改善**のループを自動で回します¹。人手による逐次的な介入無しにこのループを回せるため、**エージェントが自律的に材料空間を探索**することが可能です⁴。このような自律性により、「材料探索における人間専門家の役割の一部を代行するAI」が実現されています。

- **解釈性（説明可能な推論）**： MatAgent最大の特徴の一つが、その**高い解釈性**です。通常の生成モデルでは、「なぜその材料を提案したのか」という理由がブラックボックスになりがちですが、MatAgentではLLMが各ステップで**自然言語による理由説明**を生成します²。例えば「以前の案よりも原子半径の小さい元素を加えることで格子エネルギーが下がり安定化すると考えた」等、提案材料が有望な根拠を人間が読める文章で出力します。研究者はこの説明を読むことでAIの意思決定過程を追跡でき、結果に納得感を持てるだけでなく、新たな洞察を得ることも期待できます²。実際、本手法により**ブラックボックスだった探索過程を人間が理解可能にすることができ**、「その出力から専門家が新たな知見を獲得することも期待される」とされています²。このようにMatAgentは**説明可能AI (XAI)**の一種であり、産業応用上も重要な「AIへの信頼性」を高める効果があります。

- **専門家のような推論プロセス**： MatAgentではLLMがまるで熟練の材料研究者のように推論します。その鍵となるのが、人間の専門家が用いる**知識や戦略**を模倣した**外部ツール群の統合**です⁴。MatAgentには次のようなツールが組み込まれています：⁴ ¹

- **短期記憶 (Short-term memory)**： 直近の試行で提案した材料や得られたフィードバックを記録し、LLMが次の提案時に「これまで何を試しどうだったか」を把握できるようにします。これにより同じ失敗を繰り返さず、**逐次的な学習**が可能です¹⁴。

- **長期記憶 (Long-term memory)**： 複数の試行にまたがる知見（例えば「元素Aを加えると一般に安定性が上がりやすい」等）を保存するデータベースで、LLMはこれまでの累積知識を参照できます。時間の経過とともに**経験知**を蓄積し、より高度な戦略立案に活かします¹⁴。

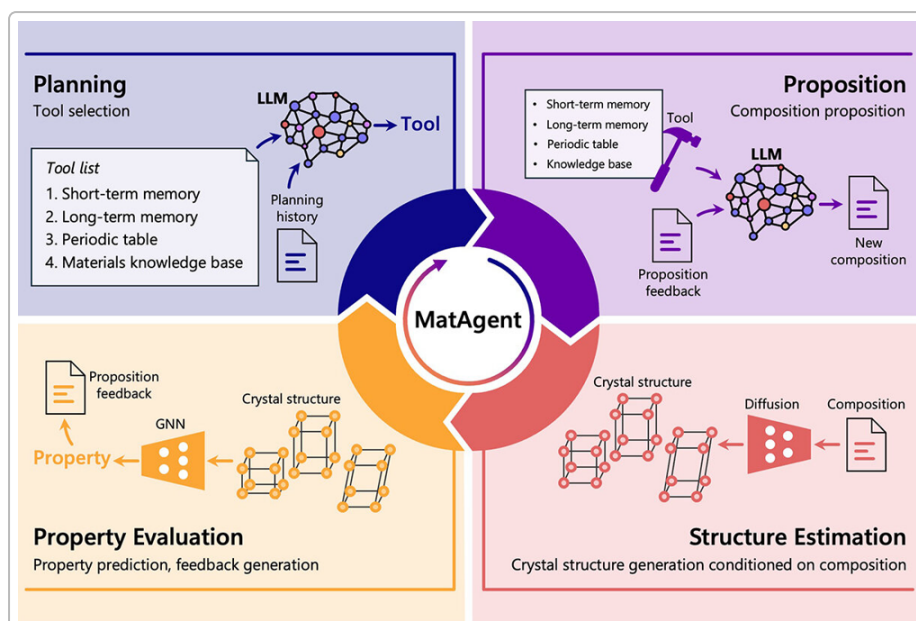
- **周期表ツール**： 周期表上の元素の性質や傾向を照会できるツールです。例えば「隣接する元素との差異」や「周期表で同族元素の性質」など、**元素の基本的な化学知識**をLLMが動的に取得できます¹⁴。専門家が暗黙知的に使う周期表知識をAIに与える仕組みです。

- **知識ベース**： 事前に構築した**材料知識データベース**で、ある組成を別の組成に変化させた際の物性変化や専門家の解説が蓄積されています¹⁵¹⁶。MatAgentでは過去の研究やデータから1万件の組成ペアとその物性変化を登録し、GPT-4による自然言語の解説文も添付されています¹⁷。LLMは現在検討中の組成に近いエントリを知識ベースから検索し（組成をベクトル化して類似度検索）、関連するエントリの解説を参照します¹⁸。これは人間の研究者が「文献を紐解く」行為に相当し、未知の組成でも**既存知識を応用した推論**を可能にします。

¹⁹ ¹ 上述のように**短期記憶**から**知識ベース**まで統合したLLMエージェントにより、MatAgentは単なるデータ駆動ではなく**知識駆動型の探索**を実現しています。研究者は「材料探索における専門家の思考プロセスを模倣したツール統合」によって、「LLMが多様な材料を提案することが可能になった」と述べています¹。さらにユーザはLLMに対し「**有害な元素を避けて**」や「**一般的に入手しやすい元素のみで**」といった制約を日本語・英語など自然言語で与えることもでき、LLMはそれを理解して探索に反映します²⁰。この柔軟性も、人間の指示を言葉で理解できる**エージェントAI**ならではの利点です。

以上の特徴により、MatAgentは**材料開発における協働的パートナー**となり得る重要な一歩だと位置付けられています²¹。材料候補の提案だけでなく、その**理由・知見まで含めて提示**できる点で、従来のブラックボックスな生成モデルとは一線を画しています。

【図1】 MatAgentフレームワーク概略図 - LLMが中心となり計画・提案を行い、構造生成と特性評価モジュールからのフィードバックで探索を洗練してゆく¹（東京大学生産技術研究所プレスリリースより）



図中の灰色枠がLLMによるPlanning/Proposition、青枠がStructure Estimator（構造生成）、緑枠がProperty Evaluator（特性評価）を示す。点線矢印はLLMが参照可能な4種の外部ツール（短期記憶=STM、長期記憶=LTM、Periodic Table、Knowledge Base）。評価結果はフィードバックとしてLLMに返される。

2. MatAgent関連論文から見るLLM・アーキテクチャ・プロンプト設計

MatAgentの詳細技術は、2025年12月に学術誌Cell Reports Physical Scienceに掲載された論文「Accelerated Inorganic Materials Design with Generative AI Agents」にまとめられています^{22 23}。この論文および対応する事前公開論文²⁴から、MatAgentに用いられたLLMの種類、システム全体のアーキテクチャ、およびプロンプトエンジニアリングの工夫について掘り下げます。

- **用いられたLLM:** MatAgentの中核となる言語モデルには、OpenAI社のGPT-4 系列モデルと、2025年初頭に公開された小型LLM「o3-mini」が採用されています²⁵。具体的には「GPT-4o（2024-08-06版）」と「o3-mini（2025-01-31版）」というモデルIDが記載されており、GPT-4をベースに科学技術分野向けに調整されたモデル（GPT-4o）と、推論性能を重視した軽量モデル（o3-mini）の2種類でフレームワークを検証しています²⁵。GPT-4oは汎用の強力なLLMで多様なタスクに適用可能な**高性能モデル**、o3-miniは**高速で科学分野に強いモデル**として位置づけられ、MatAgentでは両者の挙動や成果を比較しています^{26 27}。このように複数のLLMバックボーンを試すことで、モデル性能の違いが材料探索結果に与える影響（例えば成功率やツール使用パターンの差異）も分析しています²⁸²⁹。なお、GPT-4クラスの大規模モデルを用いることで高度な推論と自然言語理解を実現しつつ、小型モデルでもある程度の成果が出ることを示すことで、将来的な実装コスト低減の可能性も示唆しています。

- **アーキテクチャ構成:** MatAgent全体は、LLMエージェント＋外部ツール群＋生成モデル＋評価モデルから構成されています^{30 31}。その動作サイクルは**4つのステージ**に分かれています³²：

- **Planning（計画立案）ステージ:** LLMが現在の目標（ターゲット特性値）とこれまでの結果を踏まえ、次取るべき方針を言語で考えます。ここでLLMは必要に応じて**短期記憶・長期記憶・周期表・知識ベース**の各ツールから情報を取得します³³。例えば「前は安定性が不足したので、今回は安定化しやすい元素を追加する計画にしよう」「周期表から見ると隣の元素Xは電気陰性度が高い」等、**次の提案方針を論理立てて記述**します³⁴。これによりLLMが行う一連の推論に**戦略性**が与えら

れます。Planning用の詳細なプロンプトテンプレートは付録A.1に掲載されており、ツールを参照した場合は何を得たか、なぜそのツールを使ったかまで明示的に理由付けを行わせる設計になっています

35 34。

・**Proposition (組成提案) ステージ:** 続いてLLMは、具体的な**材料の組成**を新たに提案します 33。Planningで立てた方針に基づき、「どの元素を組み合わせれば目標特性に近づくか」を推論し、化学式の形で出力します。併せて「この組成を選んだ理由」も文章で説明します 33。例えば「元素AとBを組み合わせれば電子構造的に安定化し、目標の形成エネルギーに近づくと考えました」など、**提案組成が有望な根拠**を述べます 2。このProposition段階の出力（組成案と説明）は、次の評価段階で使われるだけでなく、LLM自身の短期記憶にも蓄積され次回以降のPlanningに反映されます。提案時にもツール類の情報を参照した場合、その内容を説明に織り交ぜて**説得力のある提案**になるよう促しています（付録のプロンプトテンプレートA.2～A.5に各ツール使用時のフォーマットが示されています 35 16）。

・**Structure Estimation (構造生成) ステージ:** LLMが提案した化学組成に対し、対応する**結晶構造を生成**します。MatAgentではこの役割を**拡散モデル**ベースの生成モデルが担います 30。拡散モデルはMicrosoftのMatterGenと同様に、ユニットセル内の原子配置・格子定数・対称性などを一括して生成できる手法で、本研究では既知の安定結晶データで事前学習されたモデルを使用しています 36 7。提案組成から可能な結晶構造は一意に定まらないため、Structure Estimatorはランダム性を伴って複数の候補構造をサンプルします。その中から**物性評価モデル**が最も良い特性を示す構造を選択し、次の評価ステージに渡します 37。MatAgent論文では、この部分に特化したモデルとしてZeniらによる拡散生成モデル（MatterGen） 38 や他の最新手法も参照していますが、基本的には「**提案組成**」→「**候補構造群**」という変換を担うコンポーネントです。拡散モデル導入により、組成に対応する原子配列を合理的かつ多様に探索でき、LLMに構造面の詳細を任せず**専門モデルで補完**するアーキテクチャになっています。

・**Property Evaluation (特性評価) ステージ:** 構造生成された候補の中から、目標特性に最も近いものを選び出す**評価モデル**です 37。MatAgentでは**グラフニューラルネットワーク (GNN)** ベースのモデルであるiComFormer（形成エネルギー予測に特化したTransformer）を使用し、各候補構造の**形成エネルギー (eV/atom)**を推定しています 39。生成候補の中から最も形成エネルギーが低い（安定な）構造を「選択された構造」として確定し、その組成に対する評価値（例：予測形成エネルギー）が決定されます 37。さらに、LLMにフィードバックする際には「提案した組成Xの予測形成エネルギーはY eV/atomで、目標値との差はZです」といった**フォーマット化したフィードバック文**を生成します 40 41。加えて、もしLLMの出力した組成表記が不正確・曖昧だった場合（存在しない元素記号を含む等）はエラーを検出し、「組成の形式が不適切です」といった訂正要求のフィードバックも返します 42 43。この評価結果フィードバックは次回のLLM推論（PlanningおよびProposition）に入力され、例えば「前は目標よりエネルギーが高かったので次は下げる方向で」といった反省・改良に使われます 44。

以上がMatAgentの1サイクルであり、研究ではこのサイクルを**15回程度繰り返す**ことで目標特性に合致する材料組成を見出しています 25 28。初期の組成はランダムに与える方法と、LLM自身に「まず何か提案せよ」と言う方法、また別途学習したレトリバーモデルで当たりを引く方法の3通りが検討されており、LLMに自由に考えさせても有望な初期案を出せることが示唆されています 45 46。

・**プロンプトエンジニアリング:** 上記の通り、MatAgentではLLMに多段階のタスク（計画→提案）を遂行させるため詳細なプロンプト設計がなされています。論文の付録では**各ツールを使う場合のプロンプトテンプレート**が図示され、LLM出力のフォーマットや含めるべき情報が規定されています 35 17。例えばPlanningステージ用プロンプトでは、「<目標>」「<これまでの提案>」「<フィードバック>」など変数が用意され、LLMがまず**どのツールを使うか選択**し、その上で「なぜそのツールを使うのか、得られた情報は何か」を理由付けしつつ計画を述べるよう指示されます 35 47。

Propositionステージのプロンプトも、選択したツールに応じてテンプレートが変わり、例えば知識ベースを参照した場合は「<Information>」欄に知識ベースから得た関連知見（「組成をXからYに変えると特性が〇〇変化した」というデータと解説）が与えられます¹⁷。LLMはそれら外部情報を引用しながら「ゆえに今回は元素〇を加えるべきと判断した」等と説明付きで組成提案を行います⁴⁸⁴⁹。このように**ツール使用→情報取得→判断→提案**まで一連の流れが一つの大きなプロンプトで誘導されており、LLMは定められたフォーマットに沿って応答することで、後段の構造・評価モジュールが読み取り可能な形式（きちんとした化学式や数値）を出力します⁴³⁵⁰。不適切な出力には評価ステージでフィードバックが返るため、LLMは次回以降フォーマットミスを避けるように学習できます。このプロンプト設計とフィードバック制御により、LLMの創造性と厳密さのバランスが取られ、**暴走せず目的指向に探索を進めるエージェント**が成立しています。

以上、MatAgentの関連論文から技術的裏付けを見てきました。要点をまとめると、**GPT-4クラスの強力なLLMをコアに、四つの外部ツールと生成・評価モジュールを統合し、巧みなプロンプトで制御することで、**材料空間を人間さながらに論理的かつ自律的に探索できる枠組みになっています³¹³²。このアプローチは高度にモジュール化されており、例えばより優れた生成モデルや物性予測モデルが登場すれば差し替えて性能向上させることも可能です⁵¹。実際著者らも「本フレームワークは材料科学分野の今後の生成モデル・予測モデルの発展から直接恩恵を受け、さらにそれらの発展を促進し得る」と述べており⁵¹、MatAgent自体が**オープンプラットフォーム**的に今後発展していく余地を持っています。

3. 企業主導の材料探索AI最新動向（2024～2025年）

MatAgentの登場背景には、近年急速に進む**生成AI・AI駆動型の材料探索**という研究潮流があります。特に2024年から2025年にかけて、Google・DeepMindやMicrosoft、IBMといった大手企業が相次いで大規模プロジェクトの成果を発表し、また関連スタートアップも台頭しています。本章では主要な企業の取り組みを概観し、MatAgentの位置づけを明確にします。

3.1 Google DeepMind: GNoMEによる大規模材料探索

AI企業の中でもDeepMind（現在はGoogle DeepMind）が進める材料プロジェクト「**GNoME**」は、象徴的な成果として知られます。GNoMEは**Graph Networks for Materials Exploration**の略で、グラフニューラルネットワーク技術を核にしたモデルです⁵²。2023年11月にNature誌で発表された研究では、GNoMEを用いて**220万種以上の新規無機結晶材料**の安定性を計算予測し、そのうち約**38万種が熱力学的に安定**であることを報告しました⁶。この数は、従来人類が認識してきた安定結晶（約5万種）の約10倍に相当し、「800年分の知識に匹敵する」発見だとされています⁶⁵³。GNoMEは材料の組成・結晶構造を原子間グラフで表現し、その安定性（形成エネルギーなど）を高速に予測するモデルで、巨大なデータセット上での**スケール効果**によって高精度化を実現しています⁵²⁵⁴。実際、Materials Projectや他の既存データから約4.8万種の安定構造が知られていましたが、GNoMEはそれを遥かに超える安定構造候補を提示し、新材料探索の地平を拡大しました⁵⁵⁵⁶。また外部検証として、提案された構造の一部（736種）は他の研究者によって実験的に合成・確認され、信頼性も示されています⁵⁷。DeepMindはGNoMEの予測データをオープンにしており、38万種の安定材料は米国のMaterials Projectデータベースにも組み込まれつつあります⁵⁸⁵⁹。さらにDeepMindは、ローンスバークレー国立研究所（LBNL）の自動合成ロボットと協働し、GNoMEの予測を基に実際に材料を自動合成する試みも発表しています⁵⁷。GNoMEは**網羅的かつ予測精度重視**のアプローチで、材料空間を機械学習で包括的に探索する代表例と言えます。その一方で、提案理由の説明などは主眼ではなく、どちらかといえば**データ駆動の大量探索による性能最大化**にフォーカスしたプロジェクトです。卓越した計算リソースと大規模データで可能になった成果であり、AIで材料候補を一挙にリストアップし実験者に提供するスタイルは、現在の産業界の**ハイスループット材料探索**の方向性を示しています。

3.2 Microsoft: MatterGenによる生成モデルと条件制御

Microsoftは研究部門（AI for Science）を中心に、材料設計への生成AI適用を進めています。その代表が「**MatterGen**」と呼ばれる生成モデルです⁶⁰。MatterGenは**拡散モデル**(Diffusion Model)を用いて新規材料

構造を直接創出するもので、「テキストから画像を生成するように、検索ではなく生成で材料を設計する」ことを目指しています^{61 62}。技術的には、結晶の単位胞（格子パラメータ・原子位置・元素種類という3要素）を同時に扱えるような特殊な拡散過程を設計し、**周期性や回転対称性**など結晶特有の性質も組み込んだモデルとなっています⁷。2024年6月のMicrosoft Research Forumで発表された内容によれば、まず既知の安定物質データベースで拡散モデルを学習させ、無条件でも安定構造を生成できるようにします⁸。次に、特定の**条件（ターゲット特性や含む元素など）**を与えて材料を生成するために、**ControlNet**に似た手法でモデルに条件を与えるレイヤーを追加し、少量のラベル付きデータで微調整します⁸。このファインチューニングにより、例えば「リチウムを含みイオン伝導性が高い材料」や「特定の結晶対称性を持つ材料」といった**ユーザの要求に直接応じた材料生成が可能になります**^{63 8}。MatterGenの特徴は、従来は「材料データベースを検索→絞り込み」という手順を踏んでいたのを、「初めから条件を満たす候補を生成する」ことで**探索のパラダイムを転換**する点です⁶⁴。実績として、ストロンチウム-バナジウム-酸素系で新規低エネルギー構造を生成し、ランダム探索や単純置換法より低い形成エネルギーの安定構造を見出した例が紹介されています⁶⁵。さらに磁性や機械特性など複数分野で、学習データを超える極端な物性値（例：非常に強磁性の材料）の生成にも成功し、**極限特性材料の発見**に生成モデルが有効であることを示しています⁶⁶。MatterGenの成果は2024年～2025年に論文としても公表されており、例えばZeniらのNature論文⁶⁷では安定かつ多様な無機材料を生成できる拡散モデルとしてMatterGenの詳細が報告されています。Microsoftはこの技術を社内外の材料開発に役立てるべく、今後は実験検証（提案材料の合成）を進めるとしています⁶⁸。総じてMatterGenは**生成AIで直接「当たり」を作る**アプローチであり、LLMの論理ではなく**モデルの学習能力**によって材料探索を加速する方向性と言えます。

3.3 IBM: 基盤モデルによる化学・材料探索とMolGX

IBMは材料・化学分野での**Foundation Model（基盤モデル）**戦略を掲げ、大規模な汎用モデルの構築とオープンソース化を進めています⁹。2024年末には、数億～十億規模の分子データで事前学習したTransformer系モデル群（**SMILES-TED, SELFIES-TED, MHG-GED**など）を公開しました⁶⁹。これらは分子構造をそれぞれ**文字列表現**（SMILES/SELFIES）や**グラフ表現**でエンコードし、大量データから化学の言語的・構造的パターンを学習したモデルです^{70 71}。IBMはさらに、こうした異なる表現のモデルを**Mixture-of-Experts (MoE)**というアーキテクチャで融合し、一つの問いに対して最適な「専門家」モデルの組合せが動的に選択される仕組みを開発しました¹⁰。2024年NeurIPSにて、この**マルチビューMoEモデル**がStanfordのMoleculeNetベンチマークで従来の単一モーダルモデルを上回る精度を達成したと報告しています^{72 73}。例えば毒性予測にはテキスト系表現モデルが主に働き、溶解度予測にはグラフモデルも動員されるなど、タスクに応じ自動で適材適所の特徴を活かすことが確認されています^{74 75}。このような基盤モデル群により、研究者は自前で大規模データ学習をせずとも**汎用の化学知識を持つモデル**をカスタマイズ利用でき、持続可能な材料（例えばPFAS代替や新しいバッテリー材料）探索に役立てられるとしています^{76 77}。実際IBMは、生成AIで新材料を設計するクラウドプラットフォーム**MolGX**（Molecule Generation Experience）を2021年に立ち上げ、ユーザーがターゲット特性を入力すると候補分子の構造を自動創出するサービスを提供しています^{78 79}。MolGXは社内でフォトレジスト材料の開発に適用され、人間化学者より10倍以上高速に有望分子を提案した実績もあります⁸⁰。このようにIBMのアプローチは**大規模汎用モデルとクラウドプラットフォーム**によって、多様な材料開発ニーズに応えるものです。LLMそのものを使って推論させるよりは、化学の構造と性質のマッピングをモデルが学習し設計に活かすスタイルで、企業として積み上げてきた化学情報資産や計算資源を背景に、**オープンイノベーション**も視野に入れた展開をしています^{81 82}。

3.4 その他スタートアップ等の動向

大企業のみならず、スタートアップ企業もAI駆動型の材料開発に参入しています。例えば**Kebotix**は実験室ロボットと機械学習を組み合わせる新素材発見を自動化するソリューションを提供しており、膨大な組み合わせ実験をAI制御で行う「**自律実験ラボ**」の先駆けです¹¹。また**Citrine Informatics**は材料データプラットフォームを構築し、企業が保有する材料・プロセスデータを機械学習モデルで分析して最適材料を導出する支援を行っています。電池材料スタートアップの**Mitra Chem**は、鉄系カソード材料に特化したデータとAIモデルで候補材料を絞り込み、月何千もの組成をハイスループット評価する体制を整えています⁸³。触媒分野

ではCopernic Catalystsがアンモニア合成用の新規触媒をDFTシミュレーション+MLで設計し、大手化学メーカーと試験を進めています⁸⁴⁸⁵。さらにNobleAIのように、物理モデルとMLを融合させてシミュレーションを高速化し材料開発を効率化するプラットフォームも登場しています⁸⁶。これらスタートアップの多くは、特定の産業課題（脱炭素やレアメタル代替など）にフォーカスし、**AIでニッチ課題を解決**する戦略をとっています。共通しているのは、**データ駆動と実験自動化**による開発期間短縮・コスト削減という価値提案です。LLMのような汎用推論AIを前面に押し出す例はまだ少ないものの、一部では研究論文から知識を抽出するLLMや、自社データベースに問合せ可能なチャットボットを研究者支援に使う試みもあります。スタートアップ領域では、大企業がカバーしきれない**領域特化型のAIソリューション**が次々と模索されており、材料開発におけるAI活用のエコシステムが拡大している状況です。

4. MatAgentの「専門家的」アプローチと企業アプローチの比較分析

以上を踏まえ、MatAgentに代表されるアカデミックなLLMエージェント型手法と、企業による大規模データ・計算資源を駆使した手法を対比し、その違いを考察します。特に**解釈可能性（説明可能性）**、**探索プロセス（論理的推論 vs 網羅的探索）**、**対象とする材料空間・目的の観点**から比較します。

・**解釈可能性の比較**： MatAgentはLLMによる理由説明を備えており、提案の背景や根拠を人間が理解できる形で提供します²。例えば「なぜ元素Xを加えたのか」「どの知識に基づいた判断か」を逐一説明するため、ユーザー（研究者）はAIの推論過程を追跡可能です。これは「**ブラックボックスなモデルからの脱却**」を目指すもので、研究者がAIを信頼し協働する上で大きな利点となります²。一方、企業の大規模アプローチ（GNoMEやMatterGenなど）は、その出力（新物質候補）の**妥当性は高いものの、その判断根拠はモデル内部に隠蔽**されており必ずしも透明ではありません。例えばGNoMEは数百万もの材料を予測できますが、それらの安定性に寄与する因子（電子構造上の特徴など）をモデルが明示するわけではなく、後から人間が解析する必要があります。MatterGenも生成結果は示せても、「なぜそれが極端な物性を持つと考えたか」はモデルの潜在変数に埋もれており、ユーザーにはブラックボックスです。実際、MatAgentのプレスリリースでも「既存手法では生成された材料がなぜ有望か理解しづらいブラックボックス性の課題」が指摘されており³、MatAgentはそれを克服するものと位置づけられています²。要するに、**MatAgent＝高い説明力、企業流＝結果重視で説明は二の次**という対比が見て取れます。ただし近年、企業側も説明可能性に関心を持ち始めており、例えばIBMのMoEモデルでは各タスクでどの表現が効いているか解析し化学的解釈を与える試みもあります⁷⁴。今後、企業アプローチにもMatAgent的な説明能力が取り入れられる可能性はありますが、現時点ではMatAgentが学術的探究として際立って「**解釈性**」を追求している状況です。

・**探索プロセスの比較（論理推論 vs 網羅探索）**： MatAgentは**逐次的な論理推論**でゴールに迫るアプローチです。過去の試行から学び、知識を動員し、一手一指し進めるように材料を改良していきます³⁴¹⁵。この様子はまさに人間研究者が「仮説→実験→検証→次の仮説」と試行錯誤する過程を模倣したもので、探索経路にストーリー性があります。一方、GNoMEのような手法は膨大な組成空間を**一網打尽に計算**し、安定/不安定を分類する**一括処理型**です。またMatterGenも**生成→スクリーニング**という一発勝負に近く、必要なら条件を変えて繰り返し生成しますが、基本的に各生成試行は独立です。すなわち企業アプローチは「**広く浅く全体を探り、良さそうなものを拾う**」のに対し、MatAgentは「**深く考えながら段階的に絞り込む**」違いがあります⁸⁷⁸⁸。例えば、MatAgentは15回程度の反復で目的に到達しましたが²⁵、GNoMEは2.2万種の元素組み合わせを網羅しモデル推論しています⁶。MatAgentは人間と同様に**試行回数は少なくとも効率的な道筋**を辿りますが、その道筋自体はモデル知識に依存するため局所解に陥るリスクもあります。他方、企業流は全体を俯瞰する網羅性が強みですが計算コストが莫大で、また絞り込んだ後に人間が解釈・判断する労力が必要です。論文でも「既存の生成アプローチはシングルステップでターゲット特性に合致させるのが難しく、逆に潜在空間で最適化する方法はなぜその結果になるか解釈困難だ」と指摘され⁸⁷、MatAgentの**マルチステップ逐次最適化**の有用性が強調されています⁸⁹³¹。つまり、**探索のプロセス**に関しては、MatAgentが**演繹的・計画的**であるのに対し、企業的手法は**帰納的・包括的**と言えます。どちら

が優れるかは目的によりますが、「探索過程の理解も重視するか否か」が両者を分けるポイントでしょう。

・**材料空間の対象と目的の比較**: MatAgentは現在のところ、無機結晶材料の組成とその安定性（形成エネルギー）に主眼を置いています³⁹。論文では形成エネルギーをターゲット値として、特定の値（安定～不安定の範囲）に合う組成を提案させるタスクを実証しています²⁷²⁹。将来的には複数特性の同時最適化や合成しやすさ（実現性）の評価も取り込む計画が述べられています⁹⁰。一方、企業側を見ると、DeepMind GNoMEは結晶構造の**安定性予測**がメインで、そこから派生して新奇構造や特定分野（例えば2次元材料や電池材料）の探索に応用する動きがあります⁹¹⁹²。Microsoft MatterGenは汎用的に「**所望の特性**」を条件にできるとされ、エネルギー材料・磁性材料・機械材料など広範な応用先を想定しています⁶⁵⁶⁶。IBMは分子レベルで薬剤や触媒、高分子材料など**化学全般**にアプローチしており、材料空間としては有機分子やポリマーが中心です⁷⁶⁹³。スタートアップ各社は電池正極や触媒、複合材料リサイクルなど**ドメイン特化**の問題設定が多いです⁹⁴⁸⁴。このように**ターゲット領域**を見ると、MatAgent（および学界）は**汎用フレームワークの提案**が主で具体的な実用材料創出は今後ですが、企業は**自社の関心領域に直結した性能最大化**（例えば「最高性能の電池材料を見つける」「環境対応の触媒を設計する」）がゴールに据えられています。言い換えれば、企業は**事業価値の高い目標**にフォーカスし、その達成手段としてAIを使っているのに対し、MatAgentは**AIそのものの新しい可能性**を提示し将来様々な目標に応用できる基盤技術を築こうとしていると言えます。両者の目標は重なる部分もありますが、短期的にはMatAgentは「**技術コンセプト実証**」の色合いが強く、企業側は「**既存問題のソリューション**」に直結した実務志向といった違いがあります。

以上、MatAgent流と企業流を比較すると、**解釈性・探索プロセス・目的志向**の各面で明確なコントラストが見られます。MatAgentはAIを研究者の相棒として位置づけ、人間の知的プロセスを再現・拡張する方向性であり、企業の多くはAIを強力な計算エンジンないしはデータ処理装置として活用し成果を得る方向性です。どちらが優れるというより、アプローチの**価値基準の置き所**が異なると言えるでしょう。

5. 学術的・本質的な差異: 「効率・性能最大化」vs「科学的理解・暗黙知の獲得」

企業の材料AI開発は往々にして「いかに早く・安く目標性能を達成するか」という**効率や性能指標の最大化**に重点があります。これはR&D投資に対するリターンを追求する産業界の性質上当然とも言えます。一方でMatAgentのような学術研究の文脈では、「AIが人間のように考えることで何をもたらせるか」「ブラックボックスだった部分に新たな理解を与えられないか」といった**科学的・本質的な問い**が重視されます。この節では、そのような**学術的価値の差異**について考察します。

まず企業アプローチの効率志向を見てみます。例えばDeepMindのGNoMEは、従来何十年もかかる探索をAIで一気に行い「**800年分の知識**」を一度に得ると表現されています⁶。これはまさに探索効率の飛躍的向上が主眼にあります。同様にMicrosoft MatterGenも「従来の候補探索をバイパスして、直接答えを生成する」という効率化の主張をしています⁶¹。IBMのMolGXも「新材料開発を10倍～100倍高速化する」がキャッチフレーズです⁷⁸。これらは突き詰めると「**結果が出ればプロセスは問わない**」に近い考えで、極端に言えば出力された材料が良ければそれがどう導かれたかはブラックボックスでも構わないというスタンスになります。ビジネス的には、その材料が新商品・新技術につながり利益を生めば十分であり、AIの内部原理を理解すること自体は優先度が下がります。

対してMatAgentの研究は、「AIが専門家の暗黙知を獲得し、それを共有できる形にする」という**知の拡張**を目指しています。材料科学には経験則や勘所といった暗黙知が多く存在しますが、それを人間からAIへ、さらにAIから人間へフィードバックする形で共有しようとする点に本質的価値があります。MatAgentでは知識ベースに「**組成変化による特性変化**」の**データと言語説明**を蓄積し、それをLLMが利用することで、暗黙知を再利用しています¹⁵¹⁶。さらにLLMが出力する理由説明を人間研究者が読むことで、「なるほど、その発

想はなかった」と新しい知見を得る可能性があります²。実際プレスリリースでも「出力から専門家が新たな知見を獲得することも期待される」と述べられています²。このように**AIによる暗黙知の形式知化および新たな知識創出**が、MatAgentの学術的意義です。効率やスピードよりも、AIと人間の相互作用から得られる「**理解の深化**」に重きがあります⁹⁵²。

さらに言えば、MatAgentは**AI研究そのものの進歩**にも寄与します。材料探索を題材に、エージェントAIやツール使用型LLMの先進事例を示したことで、他分野への波及効果も期待できます。企業手法がクローズドな環境で成果だけ出すのに対し、学術研究は知見をオープンに共有しコミュニティ全体の知のフロンティアを押し上げます。MatAgentの論文も付録に至るまで詳細を公開し、GPT-4など外部APIに頼りながらも再現性ある形で手法を提示しています。これは**将来的に誰もが自分の研究分野で「MatAgent流」のエージェントを作れる道を開いた**とも言えます。

総じて、企業が「**ゴール至上主義**」であるのに対し、MatAgentは「**プロセス重視**」かつ「**知の共有・継承**」を大切にするアプローチです。産業界では往々にして「結果が出ればOK」で終わりがちですが、学術的には「なぜそうなるのか」「それによって何を学べるのか」が問われます。MatAgentはまさに後者を体現しており、**科学的理解の深化と人間の知的活動のモデリング**という本質的課題に踏み込んでいる点で、企業研究との差別化が図られています。

この差異は言い換えると、AIを**道具（ツール）**とみなすか、**パートナー/研究主体**とみなすかの違いです。企業はAIを強力な道具として使い倒す発想ですが、本研究ではAIを擬似的とはいえ思考する存在とみなし、人間と共に新知見を作る**相棒**として扱っています²¹。こうした姿勢の違いが、効率vs理解という目的の差異につながっていると考えられます。

6. 産業界における位置付け：競合か補完か？

では、MatAgentのような技術は企業の研究開発に対し**競合するのか補完的に機能するのか**、その評価を行います。結論から言えば、現時点では「**補完的**」であり、将来的には一部競合し得るが共存関係を築く可能性が高いと考えられます。

まず、MatAgentは既存の企業ツールと**直接競合する段階にはない**でしょう。理由の一つは、MatAgentが示したのは枠組みの有用性であり、特定の物質を見出して即工業化するような**実用アウトプット**までは達成していない点です。例えばGNoMEは38万の安定構造データという成果を即座にデータベースに提供しましたし、MatterGenも特定材料系で既存手法を凌ぐ構造を発見しています⁶⁵。一方MatAgentは、その枠組み自体が目新しく強力であるものの、**個別分野の最適材料をもう見つけた**という段階ではありません。企業のR&D担当者から見れば、「今すぐ自社の材料開発問題を解決してくれるソリューション」ではなく、「将来的なAI活用のヒント・方向性」を示すものと映るでしょう。従って短期的には、企業は従来のデータ駆動ツールやシミュレーションを手放すことはなく、MatAgent的手法を**補助的に試行**する段階と考えられます。

補完的である具体例として、**MatAgentを企業の既存パイプラインに組み込む**可能性が挙げられます。例えば材料候補の**事前スクリーニング**にはGNoMEのような網羅モデルを使い、出てきた上位候補群についてMatAgentのLLMに「なぜこれは有望か」を解釈させる、という使い方が考えられます。こうすれば大量候補を迅速に生成しつつ、その中から人間が選ぶ際にMatAgentの説明で判断を助けてもらうという、**ハイブリッドなワークフロー**が構築できます。あるいは逆に、MatAgentに自然言語で「○○な用途に適した材料を探して」と曖昧なリクエストを与え、LLMの知識と推論である程度候補を絞り込んでもらった後、残りを高精度モデルで評価するという流れも可能です。特に**制約条件の柔軟さ**（毒性を避ける等）や**未踏領域への仮説提案**など、純粋なデータ駆動では扱いにくい部分でMatAgentは補完的な役割を果たせます。企業の研究者にとっても、MatAgentから出力される自然言語の説明はブラックボックスモデルにはない洞察を与えてくれるため、既存のシミュレーション結果に**付加情報**を付けるツールとして有用でしょう。

長期的には、MatAgentのようなエージェントが成熟することで**競合的な側面**も出てくるかもしれません。例えば、現在企業がチームを組んで行っている材料設計（試行錯誤や文献調査）を、将来MatAgentがかなり代替できるようになれば、人間専門家の役割の一部を奪うことになります。企業内ツールとして、単なるデータ解析ソフトではなく**意思決定エージェント**が導入されれば、それは従来の材料インフォマティクスツールと競合するでしょう。しかしその段階でも、人間研究者との協働関係は維持されと考えられます。MatAgentはあくまで**人間のパートナー**として設計されており²¹、完全自律AIが人間を不要にする発想とは一線を画しています。企業文化的にも、研究者がAIエージェントの説明を読んで判断し、一緒にプロジェクトを進めるといった形が現実的です。したがって、競合というより**共進化**に近い関係で、既存ツールの高度化版として受け入れられていく可能性が高いでしょう。

また、MatAgentがすぐにでも産業利用されるかどうかは、**実用上のハードル**にも左右されます。大規模LLMの利用コストや社内データへの適用制限、応答の信頼性など、解決すべき課題があります。企業によっては機密データを扱うため外部APIのGPT-4は使えないとか、リアルタイム性が要求され長い推論は許容されないといった事情もあるでしょう。その点、小型モデルやオンプレミスモデルでも動作可能であることをMatAgent論文は検証しており（o3-miniでの実験²⁵）、工夫次第で実装ハードルは下げられそうです。現状ではMatAgent自体が研究プロトタイプですが、将来企業が**自社版MatAgent**（社内知識ベースを搭載したエージェント）を構築することも十分考えられます。その時には既存ツール群とどう統合するかが鍵で、各社の戦略次第ではMatAgent思想を取り入れたハイブリッドAIプラットフォームが主流になる可能性もあります。

最後に、MatAgentは**競合ではなく新市場を創出する技術**とも言えます。つまり従来はなかった「対話しながら進める材料開発」というスタイルを生み出し、それが研究効率だけでなく研究者の創造性やナレッジマネジメントに好影響を与えるなら、企業にとっても採用する価値が十分あります。既存の高速探索ツールと単純比較してどちらが優れているかではなく、**価値の次元が異なる技術**と捉えるべきでしょう。MatAgentが示したように、AIに専門家のような推論をさせると新たな知見獲得が期待できる²という点は、企業にとっても長期的なイノベーション力強化につながります。効率一辺倒では得られないブレークスルーの種を蒔く存在として、MatAgent型AIは**補完的かつ戦略的価値**を持つと評価できます。

7. 結論 (Conclusion)

本レポートでは、東京大学生研が開発したAIエージェント「MatAgent」の技術的特徴と、同時期に台頭する企業主導の材料探索AIとの比較を行いました。MatAgentは**LLMの論理推論力と専門知識統合**により、人間さながらの材料設計プロセスを実現し、高い自律性と解釈性を備えています⁴²。一方、Google DeepMindのGNoMEやMicrosoftのMatterGen、IBMの基盤モデルなど企業のアプローチは、**大規模データ・計算資源を投入して探索効率や性能最大化を図る**もので、即応的な成果を上げています⁶⁶⁵。両者の違いを解釈性・探索手法・目的の観点で見れば、**MatAgent=解釈可能な逐次推論で知を深める、企業AI=ブラックボックス高速探索で成果を収集する**という図式が浮かび上がりました。

この違いは学術と産業の立場の差にも起因します。企業は**ROI（投資対効果）**重視でAIを道具視するのに対し、学術研究は**知の蓄積と原理解明**に重きを置きAIをパートナー視します。MatAgentはまさに後者を体現し、AIが暗黙知を獲得・共有することで人類の材料科学の理解そのものを深めるポテンシャルを示しました²。これは**効率やスピードだけでは得られない価値**であり、産業界にとっても長期的に重要な示唆を与えるものです。

産業応用の観点では、MatAgentは既存ツールと**競合というより補完**関係にあります。現状のMatAgentは概念実証段階ですが、その長所である柔軟な知識統合や説明能力は、企業の材料開発プロセスに組み込めば**意思決定支援や新規発想のブレークスルー**をもたらす可能性があります。ゆえに、企業R&D戦略としてはMatAgent型技術を注視しつつ、自社データやドメイン知識を取り入れた独自エージェントの構築を検討する価値があります。特に、AIを単なる計算エンジンではなく「**協働する研究者**」として活用する発想は、将来的な研究開発の在り方を変革し得るものです。本分析を通じて、MatAgentの登場は企業の材料探索AIに新たな方向性を示すものであり、競争力強化に向け**従来手法と併用する戦略**が有効と考えられます。材料開発にお

けるAIエージェントの台頭は始まったばかりであり、その動向を追いつつ積極的に取り入れていくことが、今後の産業界に求められるアプローチと言えるでしょう。

参考文献・情報源: 本レポートは東京大学生産技術研究所のプレスリリース⁹⁶¹ および同研究グループの論文³⁰⁹⁷を主な情報源とし、加えてGoogle DeepMind⁶・Microsoft⁷・IBM⁹各社の公式発表や業界動向に関するニュースソース⁷⁸¹¹を参照してまとめました。各出典の詳細は本文中の引用番号【】内に示しています。

¹ ² ³ ⁴ ¹² ¹³ ¹⁴ ¹⁹ ²⁰ ²¹ ⁹⁵ ⁹⁶ 【記者発表】大規模言語モデルで専門家のように材料空間を探索——自律性と解釈性を備えた無機材料設計のためのAIエージェントを開発—— - 東京大学生産技術研究所

<https://www.iis.u-tokyo.ac.jp/ja/news/4955/>

⁵ ¹⁵ ¹⁶ ¹⁷ ¹⁸ ²⁴ ²⁵ ²⁶ ²⁷ ²⁸ ²⁹ ³⁰ ³¹ ³² ³³ ³⁴ ³⁵ ³⁶ ³⁷ ³⁸ ³⁹ ⁴⁰ ⁴¹ ⁴² ⁴³ ⁴⁴ ⁴⁵ ⁴⁶ ⁴⁷

⁴⁸ ⁴⁹ ⁵⁰ ⁵¹ ⁶⁷ ⁸⁷ ⁸⁸ ⁸⁹ ⁹⁷ [2504.00741] Accelerated Inorganic Materials Design with Generative AI

Agents

<https://arxiv.labs.arxiv.org/html/2504.00741v1>

⁶ ⁵² ⁵³ ⁵⁴ ⁵⁵ ⁵⁶ ⁵⁷ ⁵⁸ ⁵⁹ ⁹¹ ⁹² Millions of new materials discovered with deep learning - Google DeepMind

<https://deepmind.google/blog/millions-of-new-materials-discovered-with-deep-learning/>

⁷ ⁸ ⁶⁰ ⁶¹ ⁶² ⁶³ ⁶⁴ ⁶⁵ ⁶⁶ ⁶⁸ MatterGen: A Generative Model for Materials Design - Microsoft Research

<https://www.microsoft.com/en-us/research/articles/mattergen-a-generative-model-for-materials-design/>

⁹ ¹⁰ ⁶⁹ ⁷⁰ ⁷¹ ⁷² ⁷³ ⁷⁴ ⁷⁵ ⁷⁶ ⁷⁷ ⁸¹ ⁸² ⁹³ IBM open sources new AI models for materials discovery - IBM Research

<https://research.ibm.com/blog/foundation-models-for-materials>

¹¹ ⁷⁸ ⁷⁹ ⁸⁰ IBM launches AI platform to discover new materials | VentureBeat

<https://venturebeat.com/ai/ibm-launches-ai-platform-to-discover-new-materials>

²² ²³ ⁹⁰ Large Language Models Unleash AI's Potential for Autonomous and Explainable Materials Discovery - Institute of Industrial Science, the University of Tokyo

<https://www.iis.u-tokyo.ac.jp/en/news/4955/>

⁸³ ⁸⁴ ⁸⁵ ⁸⁶ ⁹⁴ Five Startups Transforming Materials Discovery for Industrial Decarbonization - Net Zero Insights

<https://netzeroinsights.com/resources/material-discovery-startups/>