

技術動向レポート

データ駆動型材料開発の現在地とこれから

サイエンスソリューション部
 コンサルタント 石田 純一

人工知能(Artificial Intelligence, AI)は画像認識、機械翻訳をはじめ様々な分野に適用され成果を上げている。その1つが材料科学分野であり、これまで材料開発に要していた時間・コストを大幅に削減できるポテンシャルを秘めた技術として注目されている。本稿では、各国の研究開発プロジェクト、分析ツール、注目技術の観点から AI を用いた材料開発の現在地を俯瞰するとともに、課題や将来展望について考察を行った。

1. はじめに

(1) AI の活用と実施事例

日本企業は高い研究開発力を武器に、半導体・電池等の材料市場で強い存在感を発揮してきた。しかし中国をはじめとする諸外国との苛烈な国際競争にさらされ、近年その足元が揺らいでいる。新規材料を上市するまでには一般に長い開発期間・膨大なコストが要求されるが、国際市場で一步先んじるためには優れた材料をいち早く供給できる効率的な開発体制が必要不可欠である。そこで、材料開発プロセスに AI を導入することで開発効率の向上を目指した試みが盛んに行われている。

たとえば、AI を用いた物性予測や材料スクリーニングにより、従来よりも高速かつ大規模な材料探索を行うことが可能になっている。九州大学をはじめとした研究グループは、AI モデルを用いたプロトン伝導性材料の探索を行い、たった一回の実験で未知のプロトン伝導性材料の合成に成功したと報告している¹。また、住友化学はベイズ最適化と呼ばれる手法を合成プロセスに活用し、望ましい性能を持つ耐熱性ポリマーを合成するためのモノマー組成比を効率的に探索することに成功している²。さらに東北大

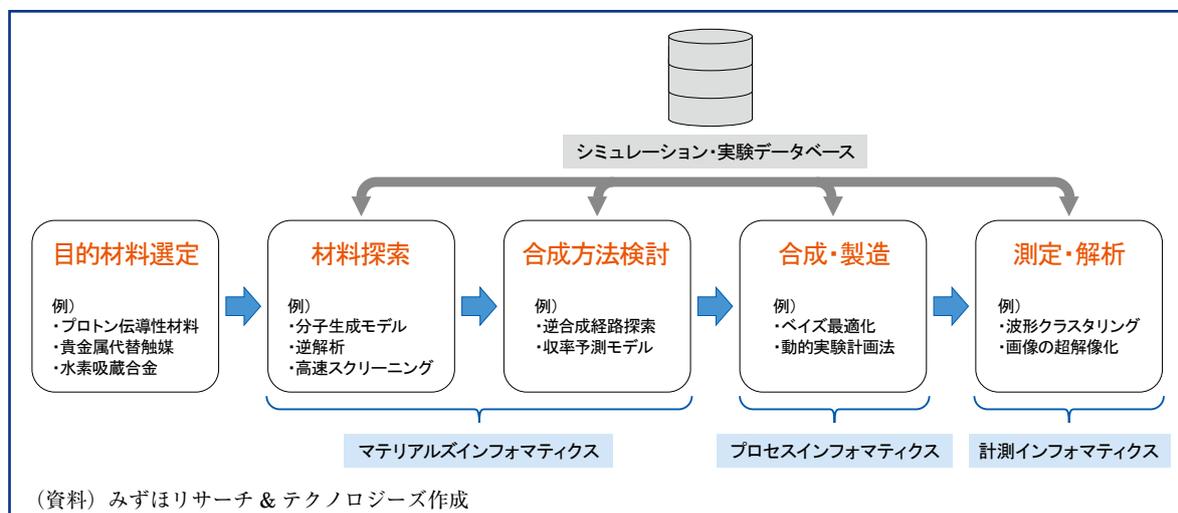
学・防衛大学の研究グループは、AI を活用することで FIB-SEM (集束イオンビーム—走査型電子顕微鏡) で取得したソフトマテリアルの3次元立体構造の超解像化や計測時間の短縮を実現している³。

このように、AI は材料開発の様々な工程・材料種に適用でき、各所で成功事例が報告されている。一方で学習データ数や材料の有効な記述子が不足しているような場合にはその真価が発揮されないこともあるため、技術の適用範囲を見極め状況に応じて AI・理論・実験・シミュレーションを適切に組み合わせることが重要となる。こうした背景から、近年では AI と材料科学を知悉した二刀流人材へのニーズが高まっており、各方面で人材育成や採用が進められている。また、後述するように豊富な計算リソースを有する IT 企業の存在感が増しており、独自の研究開発や材料系企業との提携が行われている。

(2) データ駆動型材料開発

一般に、データベースや AI を活用した材料探索など製造の前段階における手法をマテリアルズインフォマティクス (Materials Informatics, MI) と総称することが多い。また製造条件の最

図表1 データ駆動型材料開発のイメージ



適化など製造プロセスにおける AI 関連技術をプロセスインフォマティクス(Process Informatics, PI)として呼び分ける場合がある。さらに測定や解析に関する AI 関連技術は計測インフォマティクスと呼ばれることもある⁴。

本稿ではこれらの技術に明確な線引きはせず、まとめてデータ駆動型材料開発と呼ぶ(図表1)。これに関連した各国の研究開発プロジェクト、AI と研究者を結びつける分析ツール、近年の注目技術動向を整理することでデータ駆動型材料開発の現在地を明らかにするとともに、課題と今後の展望について概説する。

2. 各国における研究開発プロジェクト

(1) 米国

現在では各国がデータ駆動型材料開発を支援するプロジェクトを展開しているが、米国オバマ政権下で始まった Materials Genome Initiative (MGI, 2011-2016)はその先駆けとして知られている⁵。現在もアメリカ国立科学財団、米国エネルギー省等の各機関が MGI の派生プロジェクト(材料データベース Materials Project, 開発支援プログラム DMREF, 階層的な材料設計センター

CHiMaD など)を継続的に運用している⁶。特に Materials Project は世界中にユーザーを有する巨大データベースとして成長しており、ここに格納されているデータで学習を行った数多くの AI モデルが開発されている。諸外国の研究開発プログラムを触発するとともに、米国におけるデータ駆動型開発の基礎となる多くのプロジェクトを生み出したという点で、MGI は非常に意義深い取り組みであったと評価できる(図表2)。

(2) 欧州

欧州では The Novel Materials Discovery Laboratory (NoMaD, 2015-2018)や Materials' Revolution: Computational Design and Discovery of Novel Materials (MARVEL, 2014-2018)が先駆的なプロジェクトとして実施され、現在も後継プロジェクトが進行中である(図表2)^{7,8}。これらのプロジェクトの成果物は無償で公開されているものが多い。たとえば NOMAD Artificial Intelligence toolkit を用いることで自前の計算環境を有しない場合でも材料の AI モデリングを体験することが出来るなど、データ駆動型開発の普及に大きく貢献している。

(3)日本

日本では情報統合型物質・材料開発イニシアティブ(MI²I, 2015-2019)や超最先端材料超高速開発基盤技術プロジェクト(超超プロジェクト, 2016-2021)などのプロジェクトが行われてきた(図表2)^{9,10}。また戦略的イノベーション創造プログラム(SIP, 第一期:2014-2018, 第二期:2018-2022)により逆問題解析・マテリアルズイ

ンテグレーションシステムの開発など産官学が連携したデータ駆動型開発やそのための基盤整備が進められている¹¹。各プロジェクトのメインターゲットにはそれぞれ無機材料、有機材料、構造材料など異なる材料が選定されており、データ駆動型開発をコアとした多様な材料系での技術開発が我が国における国家プロジェクトの特徴と言える。また、2021年4月には統合イノベー

図表2 データ駆動材料開発の関連プロジェクト例

| 実施エリア | プロジェクト概要 | | 実施期間 |
|-------|---|---|------------------------|
| 米国 | Materials Genome Initiative (MGI) | | 2011-2016 |
| | 目的 | 従来の2倍の速度、1/2のコストでの材料発見・製造・展開 | |
| | 成果 | 研究開発プログラムの発足(DMREF)、材料研究開発センターの設立(CHiMaD)、材料データベースの開発(Materials Project)、MIを利用した材料開発、アウトリーチ活動など | |
| 欧州 | Materials' Revolution: Computational Design and Discovery of Novel Materials (MARVEL) | | 2014-2018 2018-2022 |
| | 目的 | MIプラットフォームを利用した新規材料の設計・発見の加速 | |
| | 成果 | 計算材料科学プラットフォーム Materials Cloud, ワークフロー管理プラットフォーム AiiDA, 計算材料科学用仮想マシン Quantum Mobile の開発など | |
| | Novel Materials Discovery (NOMAD) | | 2015-2018 2020-2023 |
| 目的 | エクサスケール規模での計算材料科学の展開 | | |
| | 成果 | 材料データプラットフォーム NOMAD Repository & Archive, 材料構造・物性データベース NOMAD Encyclopedia, AI分析ツール NOMAD Artificial Intelligence Toolkit の開発など | |
| 日本 | 情報統合型物質・材料開発イニシアティブ(MI ² I) | | 2015-2019 |
| | 目的 | MIの開拓と、アカデミア・産業界の橋渡し機能の構築 | |
| | 成果 | MI ² Iコンソーシアムによる産官学の連携、MIツールの開発(CrySPY, COMBO等)、MIを利用した磁石・蓄電池・電熱制御・熱電材料開発など | |
| | 超最先端材料超高速開発基盤技術プロジェクト(超超プロジェクト) | | 2016-2021 |
| | 目的 | 革新的な材料基盤の構築と試作回数・開発期間の1/20の短縮 | |
| | 成果 | MIを利用した有機系機能性材料開発、データプラットフォームの開発など | |
| | 戦略的イノベーション創造プログラム(SIP) 第二期:統合型材料開発システムによるマテリアル革命 | | 2018-2022 |
| | 目的 | 逆問題マテリアルズインテグレーション開発、社会実装の加速 | |
| 成果 | 産官学連携組織マテリアルズインテグレーションコンソーシアムの設立、マテリアルズインテグレーションシステムの開発、逆問題的アプローチの有効性検証など | | |

(資料) 各種情報よりみずほリサーチ & テクノロジーズ作成

ション戦略推進会議により「マテリアル革新力強化戦略」が策定され、革新的マテリアルの開発と迅速な社会実装、データ駆動型研究開発の促進、国際競争力の持続的強化がアクションプランとして掲げられており、継続的な研究開発支援が期待される¹²。

更に、近年は環境破壊や地球温暖化への危機感の高まりを受け、カーボンニュートラルの実現に向けて各国がCO₂排出量削減に関する年次目標を設定するなど具体的な取り組みを始めている。日本でも大規模な研究開発基金(グリーンイノベーション基金)が発足しており¹³、新たなエネルギーデバイス材料の創出に向けたデータ駆動型開発の試みが一層進展するものと予想される。

3. データ駆動型分析ツール

データ駆動型開発がアカデミア・民間企業各所で行われているが、一部企業においてはAIを活用した分析ツールを有償または無償で提供している(図表3)。2021年7月には、ENEOS社とPreferred Networks社が共同で開発した汎用原子レベルシミュレータ Matlantis が有償クラウドサービスとして公開された¹⁴。材料を構成す

る原子に働くポテンシャルエネルギーをAIによって予測することで、計算コストが高いDFT(Density Functional Theory)計算に比べ一万以上高速なシミュレーションを実現したとされる。従来手法に匹敵する計算精度も得られており、触媒や電池材料等の開発が大幅に効率化される可能性がある。

IBM社はAIによる物質探索の支援をサポートするため各開発工程に特化した複数のサービスを構築し、一部の機能を無償公開している^{15,16}。その一つであるIBM Molecule Generation Experienceでは、所望の物性を持つ材料をAIが自動作成する分子生成モデルをウェブブラウザ上で利用できる。また化学反応・逆合成経路探索ツール IBM RXN for Chemistry を用いて化合物の合成経路を確認することができる。

その他にも、ここ数年で国内外の多くの企業が自前のAIモデル開発やデータベース構築を行い、データ駆動型の材料分析ツールを提供している。ウェブブラウザ上での操作やクラウド環境での計算実行を特徴とするソフトウェアも多く、データ駆動型開発を行うためのハードルが徐々に低下している。

図表3 代表的なMI・AI手法と関連したデータ駆動型分析ツールの例

| MI・AI手法 | 適用開発工程 | 対応ツール例 | 開発企業 |
|-------------|--------|---------------------------------------|-----------------------------|
| エネルギー・力予測 | 材料探索 | Matlantis | ENEOS Preferred Networks |
| 物性予測 逆解析 | 材料探索 | 材料開発 ソリューション | 日立製作所 |
| 生成モデル | 材料探索 | IBM Molecule Generation Experience | IBM |
| ナレッジ グラフ | 材料探索 | TABRASA | IBM 長瀬産業 |
| 逆合成経路探索 | 合成方法検討 | IBM RXN for Chemistry | IBM |

(資料) 各種情報よりみずほリサーチ & テクノロジーズ作成

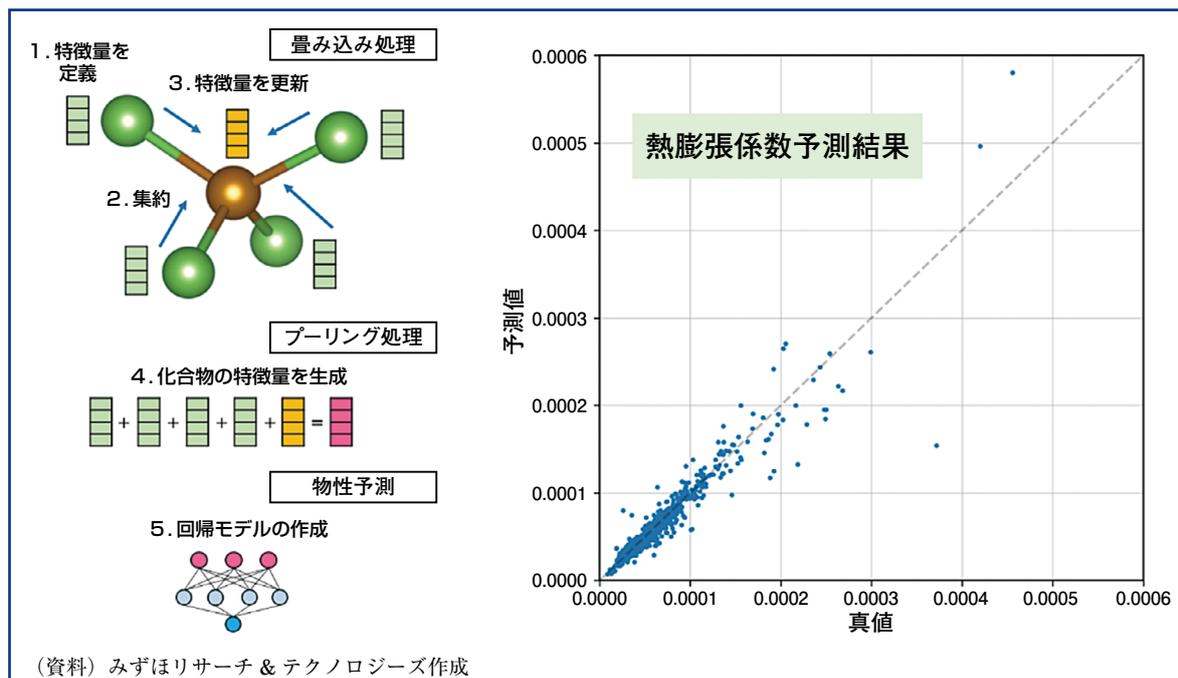
4. 注目技術動向

基礎研究レベルでも新たな技術開発は活発である。数多くの注目技術が存在するが、例えば2010年代後半に登場したグラフ畳み込みニューラルネットワーク(Graph Convolutional Neural Networks, GCNN)と呼ばれる機械学習モデルはMI手法の開発に大きなインパクトをもたらした。このモデルでは、化合物中の原子同士の結びつきをグラフ構造と見立て、グラフの頂点(原子に対応)・辺(化学結合に対応)に周辺環境を織り込んだ情報を集積し、物性予測性能を獲得している(図表4)¹⁷。さらに近年、座標情報や並進・回転対称性といった化合物の物理的性質を学習に反映させた幾何学的深層学習が注目されており、物性予測のみならずエネルギー・力の予測でも高い性能を示している¹⁸。このAIモデルを分子動力学計算に活用すると、計算コストのかかる第一原理計算をAIで置き換えた

高速・高精度シミュレーションが実現する可能性があるため、基礎研究でありながら応用側にも近い技術として非常に興味深い。

アカデミア・材料系企業のみならず巨大IT企業も材料開発を推進している。2020年にはFacebook社とCarnegie Mellon大学が共同で開発した触媒材料に関するデータベースプロジェクトOpen Catalyst Projectが公開された¹⁹。触媒材料の構造を入力として、原子毎の力や位置、構造のエネルギーを予測するための第一原理計算由来の膨大なデータが用意されている。2021年7月にはMicrosoft社が材料科学分野の研究機関をアムステルダムに新設し、機械学習を活用した分子シミュレーションの推進を標榜している²⁰。また上述のPreferred Networks社が材料・創薬分野での研究開発を進めるなど、従来はアカデミアや材料系企業の独壇場であった材料開発に国内外のIT企業が参入を進めており、今後勢力図が書き換えられる可能性がある。

図表4 GCNNの処理イメージと無機材料の熱膨張係数予測結果¹⁷



5. 現状の課題

現在ではデータ駆動型材料開発の実力が認められつつあるが、その活用にはAIそのものの性質に起因した以下のような課題が存在する。

第一に、データが少ない材料系はAIの恩恵を受けづらい。大規模系などシミュレーションによる物性予測が難しい場合や、秘密データ・実験失敗データなどオープンデータが活用しづらい場合には、少量データに基づいたモデルを構築せざるを得ず十分な精度が得られない可能性がある。

第二に、データベースや計算環境の整備に向けた設備投資や人材確保が必要となる。特に自前で第一原理計算データベースを構築する場合や、大規模データに基づく学習を行う場合にはオンプレミス・クラウドいずれの計算環境であっても相応のコストを要することになる。また、それらを維持・管理しデータを材料開発に活用できる人材がいなければ宝の持ち腐れとなるリスクがある。

第三に、AIはデータの内挿性に優れるが外挿性に欠ける。たとえば、特定の原子から構成される化合物のデータセットを用いて学習や物性予測を行った場合、それらの原子を含むテスト用化合物に対しては高い予測精度が期待できる。一方でデータセットに含まれない原子を含む化合物に対しては性能の悪化が懸念される。

6. 今後の展望

第一の課題を解決するため、事前に取得可能な大規模データで学習したモデルを小規模データに転用し再学習する技術(転移学習)が採用されることがある。近年は膨大な数のパラメータから成る汎用性の高い学習済みモデルを、各ユーザーが自身の用途に合わせてチューニングする流

れが特に自然言語処理分野で一般化しつつある。材料科学分野でも、たとえば学習済みの力場モデルをケースごとにチューニングするなどの動きが起こる可能性は否定できない。また、データ拡張・データ生成等による精度向上技術は今後も着実に進展するものと見込まれる。

第二の課題に関しては、設備投資に先立ちデータ駆動型材料開発を推進するための効果的なITソリューションを提案できる人材の確保が重要となるだろう。そのため各研究機関における人材育成や、材料系企業と豊富な計算資源・優れたAI技術を有するIT企業との合従連衡が今後のトレンドとして加速する可能性がある。

第三の課題には、ガウス過程回帰といった誤差評価が可能な予測モデルの適用や化合物情報の追加によるデータベースの拡充が有効であると思われる。また、類似データの存在しない領域の材料探索に理論・実験・シミュレーションが威力を発揮することは、これまでの材料科学の歴史が証明している通りである。そのため、未開の材料を探索する上では従来技術とAIを高度に組み合わせた技術開発が一層重要になるだろう。

AIを用いた材料開発は基礎研究レベルでは数多くの成果を生み出しているが、実際の開発現場へ適用できる技術は決して多くはなく、今後も模索が必要であると筆者は考えている。しかしながら、世界的に競争が激化する材料開発において日本が一步先をリードするためには、AI・データ・人材への果敢な投資と、「データ中心の材料開発」に向けた本格的な舵取りが必要になると思われる。

参考文献

1. 国立研究開発法人 科学技術振興機構 プレスリリース <https://www.jst.go.jp/pr/announce/20210804/index.html>
2. 日経クロステック <https://xtech.nikkei.com/atcl/>

- nxt/column/18/00026/00043/?P=2
3. 東北大学 プレスリリース https://www.tohoku.ac.jp/japanese/tohokuuniv-press20180412_deeplearning01.pdf
 4. 科学技術未来戦略ワークショップ報告書 材料創成技術を革新するプロセスインフォマティクス <https://www.jst.go.jp/crds/pdf/2020/WR/CRDS-FY2020-WR-12.pdf>
 5. Materials Genome Initiative <https://www.mgi.gov/>
 6. Materials Genome Initiative The First Five Years of the Materials Genome Initiative: Accomplishments and Technical Highlights <https://www.mgi.gov/sites/default/files/documents/mgi-accomplishments-at-5-years-august-2016.pdf>
 7. NOMAD Centre of Excellence <https://www.nomad-coe.eu/>
 8. MARVEL <https://nccr-marvel.ch/>
 9. 情報統合型物質・材料開発イニシアティブ <https://www.nims.go.jp/MII-I/>
 10. 超最先端超高速開発基盤技術プロジェクト https://www.nedo.go.jp/activities/ZZJP_100119.html
 11. 内閣府 戦略的イノベーション創造プログラム <https://www8.cao.go.jp/cstp/gaiyo/sip/>
 12. 内閣府 マテリアル戦略 <https://www8.cao.go.jp/cstp/material/material.html>
 13. グリーンイノベーション基金事業 <https://www.nedo.go.jp/activities/green-innovation.html>
 14. Preferred Computational Chemistry MATLANTIS <https://matlantis.com/ja/>
 15. IBM Molecule Generation Experience(MolGX) <https://www.ibm.com/blogs/think/jp-ja/ibm-molecule-generation-experience/>
 16. IBM RXN for Chemistry <https://rxn.res.ibm.com/>
 17. 石田純一「結晶畳み込みニューラルネットワークによる熱膨張係数予測」, みずほリサーチ & テクノロジー技報、Vol1, No1 (2021年)
 18. Shuaibi, Muhammed, *et al.* 「Rotation Invariant Graph Neural Networks using Spin Convolutions」, arXiv preprint, arXiv:2106.09575 (2021年)
 19. Facebook AI Open Catalyst Project <https://opencatalystproject.org/>
 20. Microsoft Research Lab – Amsterdam <https://www.microsoft.com/en-us/research/lab/microsoft-research-amsterdam/>