

2.5.3 データ駆動型物質・材料開発

(1) 研究開発領域の定義

4つの科学（実験科学、理論科学、計算科学、データ科学）を統合的に活用して、新物質・新材料開発を効果的に推進する研究開発領域である。新規物質・材料の設計・探索・発見を飛躍的に加速するマテリアルズ・インフォマティクスが一定の成果を示しており、今後は材料製造プロセスを最適化するプロセス・インフォマティクスや、計測・解析を効率化する計測インフォマティクスとの連携の重要性が増している。

また、それらを支えるロボットによるハイスループット実験や、AI技術を活用した自律的最適化（Closed-Loop）実験など実験DXも重要な技術要素である。

(2) キーワード

データ駆動、マテリアルズ・インフォマティクス、プロセス・インフォマティクス、計測インフォマティクス、機械学習、データ科学、ハイスループット実験、ロボティクス、自動化、自律的最適化実験、Closed-Loop、AI技術

(3) 研究開発領域の概要

[本領域の意義]

SDGsの達成やSociety 5.0の実現に向けて、材料開発に対する要求がますます高度化している。このため、材料開発は多元素化・複合化や準安定相・ハイエントロピー相の利用などの方向に向かっており、材料探索空間は急拡大している。この広大な材料探索空間を従来の材料探索技術（実験科学・計算科学・理論科学を駆使した試行錯誤的アプローチ）のみで探索するのはもはや不可能である。そのため、データ科学とAI技術を駆使したデータ駆動型物質・材料開発が必須の技術領域となり、世界的に急ピッチで技術開発が進められている。

2011年に米国が「Materials Genome Initiative (MGI)」を発表して以来、マテリアルズ・インフォマティクスは、特に、新規物質・材料の設計・探索において有力な手法であることがさまざまな研究事例をもって実証され、材料研究にデータ科学を活用することが今では当たり前となりつつある。しかし、マテリアルズ・インフォマティクスによって設計された新規物質・材料が、その合成方法が見つからずにその性能を実証できない例も散見されている。このため、今後は、材料製造プロセスを探索・最適化するプロセス・インフォマティクスや、計測・解析を効率化する計測インフォマティクスとの連携の重要性が増している。

また、広大な材料探索空間においてデータ駆動型物質・材料開発を行うために、実験DX（Digital Transformation）の活用も重要である。ロボットなどによるハイスループット実験や、AI技術によるスペクトル解釈・画像解析、これらを統合した自律的最適化手法が、材料研究分野における日本の国際競争力の強化に大きく貢献すると期待される。

[研究開発の動向]

データ駆動型物質・材料開発技術の動向を、まず、マテリアルズ・インフォマティクス、プロセス・インフォマティクス、計測インフォマティクス分けて記載する。

その後、日本、米国、欧州における主な動向について記載する。

・マテリアルズ・インフォマティクス

マテリアルズ・インフォマティクスは、広義にはデータ駆動型材料科学の全体を示すこともあるが、2010年代前半頃から特に活発に研究開発が行われてきたのは、材料特性の予測や新材料を設計する技術としての

マテリアルズ・インフォマティクス、いわば狭義のマテリアルズ・インフォマティクスであり、現在では、様々な記述子データ（数値・スペクトル・画像・グラフ・テキストなど）から、物性値の予測や新材料の組成・構造の推定が可能となっている。

データを活用して材料特性予測や新材料設計をするさまざまな技術は確立されつつあり、データが十分に揃っている領域であれば物性予測・新材料設計は可能だといえる。しかし、材料データが十分に揃っている領域はまだ限定的であるため、ハイスループット実験や大規模計算によってデータ蓄積を強化することが引き続き課題である。

• プロセス・インフォマティクス

マテリアルズ・インフォマティクスによって物性予測や新材料設計は可能となりつつあるが、実際にそれを合成するには大きな障壁がある。合成される材料の品質・性能は、その製造プロセス（合成温度、圧力、時間など）に大きく影響を受けるため、これらプロセス・パラメータを効率的に最適化する必要があるが、合成には大きなコストがかかるため、より少ない試行回数で最適化する技術が求められる。データ科学を活用して効率的に材料合成プロセスを探索する技術はプロセス・インフォマティクスと呼ばれ、近年活発に研究が進められている。

プロセス・インフォマティクスで使われるプロセス・パラメータ（合成温度、圧力、時間など）は、マテリアルズ・インフォマティクスで使われる材料パラメータ（物性値、組成、構造など）に比べてその種類が格段に多いが、このようなデータを十分に蓄積しているデータベースは非常に少ない。また、データ科学を有効に活用するためには、失敗データ（例えば、目的の材料が合成できなかった時のプロセスデータなど）が必要であるが、このようなデータはほとんど保存・蓄積・共有されていない。そのため、プロセス・インフォマティクスは、マテリアルズ・インフォマティクスよりさらに高難易度であると考えられている。

そこで、プロセスデータを効率的に蓄積するための技術として、実験機器のIoT化や、文献からテキストマイニングを活用してプロセス情報を抽出する技術、データ科学とロボティクス活用することによる実験の自動化・自律化に関する研究が活発に進められている。

• 計測インフォマティクス

計測インフォマティクスは、データ科学によって計測・解析を効率化する技術である。

特に近年、ハイスループット実験が盛んにおこなわれるようになり、人間ではデータ処理できないほどの大量の計測データが手に入るようになった。例えば放射光施設では、二次元スペクトルマッピングなどのデータが手軽に取得できるようになっているが、大量データを効率的に解析するためにデータ科学が積極的に活用されはじめています。さらに超解像技術を使うことによってこれまで計測できなかったものの計測が可能になっている事例もある。

また、データ科学とロボティクスと計測技術を組み合わせた計測の自動化・自律化についても活発に研究が進められている。

• 日本

第6期科学技術・イノベーション基本計画（2021年3月発表）において、知のフロンティアを開拓し価値創造の源泉となる研究力の強化が必要だとされ、そのための施策の一つとして、オープンサイエンスやデータ駆動型研究等の新たな研究システムの構築が挙げられている。具体的な施策として、研究データの管理・利活用、スマートラボ・AI等を活用した研究の加速や、研究施設・設備・機器の整備・共用、研究DXが開拓する新しい研究コミュニティ・環境の醸成が取り上げられている。

2021年4月には、マテリアル革新力強化戦略が発表された。マテリアル・イノベーションを創出する力である「マテリアル革新力」を強化するためのアクション・プランの1つとして「マテリアル・データと製造技術を

活用したデータ駆動型研究開発の促進」が示されており、具体的には、良質なマテリアルの実データ・ノウハウ・未利用データの収集・蓄積・利活用促進や、製造技術とデータサイエンスの融合、革新的製造プロセス技術の開発が記載されている。

マテリアル革新力強化戦略を踏まえ、文部科学省では、「マテリアルDX(デジタルトランスフォーメーション)プラットフォーム」を構築し、データ創出から、データ統合・管理、データ利活用まで一気通貫した研究DXを推進することを目指している。このために、①データ中核拠点整備、②データ創出基盤構築、③データ創出・活用型マテリアル研究開発プロジェクト開始をし、それぞれが強く連携してマテリアル研究開発を推進することを目指している。

データ中核拠点は、世界最大級の材料・物質データベースであるMatNaviを保有・運用している物資・材料研究機構(NIMS)が担い、マテリアル先端リサーチインフラ事業で新たに創出されるデータを含めて利活用できる材料データプラットフォーム事業を推進する。

データ創出基盤としては、マテリアル先端リサーチインフラ事業(ARIM)が、7つの重要なマテリアル分野をカバーする形で、全国25の大学・研究機関が参加して2021年度よりスタートしている。各機関が先端共用設備を保有し、それを利用して高品質なデータを創出する基盤としての役割が期待されている。また、得られたデータを収集・構造化し、利用しやすい形で登録するための手法の検討も進めている。また、ARIMで得られたデータは、当該研究当事者だけが利用できるクローズド方式だけでなく、ある一定の範囲の者が利活用できる方式(データシェア)にすることが望ましいが、それを実現するためには制度・仕組みの整備が重要であり、それらの検討も進めている。

データ創出・活用型マテリアル研究開発プロジェクトは、データ活用型研究を取り入れた次世代型研究方法論を実践し、革新的機能を有するマテリアル創出を目的としており、2021年度のフェージビリティースタディーを踏まえて、2022年度からは5拠点を研究開発プロジェクトが開始された。具体的には、「極限環境対応構造材料研究拠点(代表機関:東北大学)」「バイオ・高分子ビッグデータ駆動による完全循環型バイオアダプティブ材料の創出拠点(代表機関:京都大学)」「智慧とデータが拓くエレクトロニクス新材料開発拠点(代表機関:東京工業大学)」「再生可能エネルギー最大導入に向けた電気化学材料研究拠点(代表機関:東京大学)」「データ創出・活用型磁性材料開発拠点(代表機関:物質・材料研究機構)」である。

NEDO「超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年度)は、主に有機系材料を対象とし、計算・プロセス・計測の三位一体による革新的な材料基盤の構築、従来と比較して試作回数・開発期間の1/20への短縮、国内素材産業の優位性確保のためプロジェクト成果の実用化を最終目標としていた。

産業技術総合研究所では、超超プロジェクトの成果を活かし、データに基づく材料開発への変革をさらに推し進めることを目指して「データ駆動型材料設計技術利用推進コンソーシアム」を2022年4月に設立した。また、産業技術総合研究所では、マテリアル・プロセス・イノベーション・プラットフォームの構築も進めている。製造プロセスの高度化をデータ駆動型の観点から支援するために、セラミックス・合金拠点(中部センター)、先進触媒拠点(つくばセンター)、有機・バイオ材料拠点(中国センター)に、それぞれの材料の製造プロセスデータを収集し活用するための基盤整備を目指している。

・米国

MGIの初期5年間で一度終了したのは国家政策上の後継は顕在化していなかったが、NISTや主要大学では活発な活動を継続していた。2020年6月に大統領科学技術諮問会議(PCAST)が発出したレポートにおいて、ポストMGIとしての方向性が再提起され、その後、科学技術政策局(OSTP)にMaterials Genome Initiative(MGI)に関するサブコミッティが設けられ、2021年11月には、この先5年間で想定した戦略目標 MGI Strategic Plan 2021が発表された。ここでは、(1)材料イノベーション基盤(MII: Materials Innovation Infrastructure)を統合すること、(2)材料データの力を活用すること、(3)材料

研究開発の労働力について教育と訓練を行い繋げていくことの3つのゴールが掲げられている。

MIIとは、シームレスに統合された先進モデリングツール及び計算ツール、実験ツール、定量データの動的かつ発展的でアクセス可能な枠組みを指し、それらを統合することで、個々のツールの価値を高め、より簡単にアクセスできるようにし、材料開発全体にわたり全てのステークホルダーが容易に理解を共有するプラットフォームの構築を指す。また、MGIコミュニティ全体の現在及び未来のニーズに対処するために、National Materials Data Network (NMDN) の構築を掲げている。

MIIに関しては、データ流通のために満たすべきFAIR原則（Findable：データを見つけられるようにする、Accesible：アクセスできるようにする、Interoperable：相互運用できるようにする、Reusable：再利用できるようにする）に沿ったデータ管理を実現することを目指す内容である。とくに、米国が重視するAI研究との融合による産業競争力強化のためにも、このようなデータ管理が極めて重要であることを強調している。

MIIを実現するための取り組み事例として、NSFがファンディングする拠点形成のプログラムであるMaterials Innovation Platforms (MIP) が挙げられており、現在、4拠点が採択されており、2020年から5年前後の予定で取り組みが進められている。

MNDNを後押しする具体的な施策としては、オープン・アクセス可能・相互運用可能な材料データを実現することを目的として2019年11月に設立された材料研究データアライアンス (MaRDA) が挙げられている。パブリックとプライベートの両セクタのデータを統合できるようなフレームワーク開発などが期待されている。

データ・インフラストラクチャーに関しては、米国におけるデータプラットフォームシステムのいくつかにおいてデータ公開のサービス運用が開始されている。

Materials Data Facility (MDF) は、ノースウエスタン大学のCenter for Hierarchical Materials Design (CHiMaD) のデータツールの一つとして開発が進められてきたものであり、登録されるデータにDOIを付与し、グローバルな学術コミュニケーションの中でデータを流通させることができるほか、組織、登録年、キーワード、関連する元素などに注目した検索機能を有している。600件強のデータセット（2022年8月時点）が収録されている。

また、ミシガン大学のPRISMSセンターが開発したMaterials Commonsも、同様のデータ公開機能を有し、100件弱のデータセット（2022年8月時点）が収録されている。

MDFもMaterials Commonsも、Globusと呼ばれる学術データ流通サービスと連携が取れるように設計されている。データにアクセスするためにはアカウント認証が必要であるが、ORCIDと呼ばれるグローバルな研究者IDを利用してログインすることが可能である。FAIR原則を実現していく上で、データにまつわる行為とその主体者を記録することが求められるが、このような仕組みは今後ますます重要性を増すものと考えられる。

NISTでは2017年からHigh-Throughput Experimental Materials Collaboratory (HTE-MC) という取り組みが始まっており、ここでは米国内の研究施設をネットワーク化し、データ・試料を流通させて、設計、製造、評価を最適なラボで行う一貫した研究開発基盤提供を目指している。

• 欧州

ドイツ・オランダを中心とする研究機関連携の母体としてFAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V. (FAIR-DI) が2018年9月に発足している。この団体はNOMAD-CoEを軸におくが、材料分野に止まらずバイオ分野・天文学分野とも横串を指す形で、FAIR原則に従うデータ管理の実現と、そのための世界的なデータ・インフラストラクチャーを構築することを目的としている。生データを登録・管理するリポジトリサービス、解析ツール、可視化、知識ベース、メタデータの統合の合計5つのサービスを柱として発展している。2020年3月に3年間の追加ファンディングが決定し、活動を継続している。

Materials design at the eXascale (MAX) a European centre of excellenceは、エクサスケールの

High Performance Computing (HPC) のためのインフラストラクチャーである。このセンターでは、5箇所のHPCリソースをネットワークし、それらの上で動くcode開発をテクノロジーパートナーと連携して活動している。

MAXのデータ管理プラットフォームとして機能するのが、Materials Cloudである。PythonをベースとするオープンなシステムであるAiiDAをベースとして動作するもので、特に、計算データの生成ワークフローを管理し、これにともなうデータ来歴 (provenance) の管理も行えることが特色となっている。AiiDA、Materials CloudのいずれもスイスEPFLにおいて開発が進められてきたものであり、AiiDAは2022年4月のバージョンアップでストレージとの連携機能が大幅に強化されている。

European Materials Modeling Council (EMMC) は、Horizon2020ファンディングを得てデータ標準化を進めてきた。EMMCは、材料分野において産業界、とくに中小企業をカスタマーとするデータ市場の立ち上げを目標としており、市場でデータを流通させるために必要な共通記述様式の策定を進めてきた。2018年にロードマップをまとめ上げて一旦終了したが、2019年7月に非営利法人化してその後も活動を継続している。データ市場実現に向けたプロジェクトが2つあり、いずれもMODAと呼ばれる材料モデリングにおけるデータ記述のための標準規格に準拠した標準オントロジーを基礎においてデータ構造化を進めている。

一方、入出力するデータの形式を揃える別の手段として、入出力結果へのアクセスを全てAPI化し、そのAPIの仕様を標準化するという方法がある。この方式のもと、コミュニティ主導で進められている活動がOPTIMADEである。年1回ワークショップで材料分野におけるデータ入出力をREST API化することについて議論を進めており、現在、バージョン1.0.0として仕様が公開されている。欧州のNOMAD、Materials Cloudに加えて、米国の代表的な第一原理計算データベースを運営するAFLOW、Materials Project、OQMDはいずれもこの取り組みに関与しており、現在16のAPIプロバイダーがOPTIMADEに準拠しているとしている。今後、第一原理計算コミュニティにおいて、OPTIMADEはデータ入出力のデファクト標準になる可能性があり、注視する必要がある。

また、2022年3月、EUは循環型経済行動計画としてSustainable Products Initiative (SPI) を発表した。この中でDigital Product Passport (DPP) という考え方が導入された。これは製品の製造元、使用材料、リサイクル性、解体方法などの情報を備えた電子証明書で、製品のトレーサビリティを確保することでサーキュラーエコノミーを実現しようという試みであるが、今後、この多くの材料情報が含まれることが予想されており、新たな材料データ源としての活用も想定される。

• 中国

2014年に上海市と上海大学が共同で進めるShanghai Materials Genome Instituteを設立し、2016年には中国科学院物理研究所と北京科技大学が共同で北京マテリアルズゲノム工学イノベーション連盟を設立した。さらに同年、上海交通大学においてもマテリアルズゲノム連合研究センターを設立するなど、国を挙げてマテリアルゲノム研究に力を入れている。2016年3月に発表された科学技術イノベーション第13次五カ年計画においては、中国産業の国際競争力向上のための重点技術の1つ「新素材技術」のなかに「マテリアルズゲノム工学 (目標：新材料の開発期間・コストの半減)」と明記され、さらに、国家重点研究開発計画の1つとして「材料ゲノム工学のキーテクノロジーと支援プラットフォーム」(2016~2018年の3年間に計44課題、総額約8億元) が推進された。2021年3月に発表された科学技術イノベーション第14次五カ年計画では、材料ゲノム工学についての明示的な記載は見当たらないようだが、データ科学を活用した材料開発研究が引き続き推進されているものと見られる。

• 韓国

2015年から10年計画でCreative Materials Discovery Projectが実施されている。また、最近、韓国科学技術研究所 (Korea Institute of Science and Technology : KIST) において計算科学を中心とした

Materials Informatics Database for Advanced Search (MIDAS) が設置された。

(4) 注目動向

[新展開・技術トピックス]

• データ科学とロボティクスの組み合わせによる自律的最適化実験システム

一杉、清水らのグループ（東大、東工大）は、無機薄膜の自動成膜・自動物性評価とベイズ最適化を組み合わせたシステムを開発している。中央に試料搬送用ロボットアームを配置し、その周辺に自動成膜装置、自動評価装置がサテライト状に配置されている。これらの成膜・評価・最適化をコンピュータにより一括制御することで、全自動で物質探索ができるシステムとなっている。全固体電解質の成膜プロセスの最適化実験では、人間が実験する場合の約10倍のスループットが得られたとしている。

欧州については、自律実験に関する大型の研究プロジェクトに積極的である。特に英国のCroninらのグループ（Glasgow大）とCooperらのグループ（Liverpool大）の2グループが先行している。Croninらのグループは、標準的な有機合成プロセスで使用される丸底フラスコ、フィルター、分液ロート、ロータリーエバポレーターなどの従来型装置を組み合わせて、さまざまな化合物の合成に対応できる自動実験プラットフォーム「Chemputer」を構築し、代表的な医薬品化合物の自動合成が可能であることを実証している。Cooperらのグループは、一般的な自走式のロボットを改良し、実験室の中を移動し人間の化学者と同じ機器を用いた実験ができるようにしている。このロボットにより光触媒の性能向上を目的とした実験を行い、人間が行えば数ヶ月かかる検討を8日間で完了したことを報告している。このロボットシステムは従来の実験室に改良を加えることなく使えるもので、適用範囲が広いことを特徴としている。

また、Horizon EuropeのBattery Partnershipの2022年の公募（TOPIC ID: HORIZON-CL5-2022-D2-01-03）においては、持続可能なバッテリーについての、AI、ビッグデータ、全自動ロボット合成、ハイスループット評価を用いた材料開発プラットフォーム開発に大きな予算が割かれており、今後、バッテリー分野における競争が激化することが想像される。ドイツはEUが導入したDPPをEV向けバッテリーに対して先行的に適用することを発表している。

カナダでは、Berlinguetteらのグループ（British Columbia大）の進展が目覚ましい。有機薄膜の原料調整、成膜、特性評価を自律的に行うことができるモジュール式で自動運転プラットフォーム「Ada」を活用したペロブスカイト太陽電池薄膜の最適化実験では、従来は9か月かかっていたものが5日間でできるとしている。

米国においても、MIPの採択課題に“automated”、“accelerated”などの含まれるものが見られ自動化プラットフォームづくりを目指している様子が窺える。また、個別の自律実験の報告件数が多いのは米国であり、要素技術の水面下での開発が進んでいると思われる。

ただし、例えば、ロボットやAIは予め決められたプロトコルは粛々とこなすが、いつ自律実験を「終了すべきか」の客観的な指標がなく、現状では人間の主観に依存していることが多い。最近は、このような収束判定の自己判断を行う機械学習の報告も出始めている。

• 機械学習モデルの解釈性と物性物理のメカニズム理解

機械学習モデルの『解釈性（Interpretability）』も注目されている。深層学習のようなブラックボックス型の機械学習モデルは、高い予測性能を誇る一方、モデル内部の詳細な情報を抽出することが難しい。そこで近年、高い予測性能を保ちつつ高いモデル解釈性を持つ機械学習手法の開発が進められている。機械学習モデルの解釈性に関しては、材料研究領域だけではなく機械学習(AI)研究領域全体においてもホットなトピックであり、“Interpretable machine learning”や“Explainable AI (XAI)”といったカテゴリとして急速に研究が進められている。解釈性の高い機械学習を用いると、物理、化学、材料学などの知見をベースとして人間が機械学習モデル内部を解析することができるため、材料研究を推進するためのキーテクノロジーの一

つとして考えられる。

また、多量のデータから新規材料・物性を“予測”するのではなく、材料物性のメカニズムを“理解”しようと研究も進められている。ここでは、深層学習のようなブラックボックス型の機械学習だけではなく、説明性や解釈性の高いホワイトボックス型の機械学習も用いられる。機械学習モデルの説明性や解釈性に関しては、材料科学に限らず様々な分野でも注目されるホットトピックであり、世界中で研究が進められている。

[注目すべき国内外のプロジェクト]

【国内】

- ・データ創出・活用型マテリアル研究開発プロジェクト事業（2022年～）
共通的なデータ収集・蓄積・流通・利活用のための基盤整備を進めるとともに、先端共用施設・設備からのデータ創出や重要技術・実装領域を対象とする、データを活用した研究開発プロジェクト（前掲）
- ・JST-CREST / 社会課題解決を志向した革新的計測・解析システムの創出（2022年～）
計測技術の進化と最先端の数値モデリング・機械学習等の情報技術とを組み合わせ、計測・解析手法を高度に進化させることにより、現実の様々な難課題を解決でき、また、今後長期間にわたり計測・解析プロセスを革新できる、新たな計測・解析システムの創出を目指す。
- ・JST-CREST / 未踏探索空間における革新的物質の開発（2021年～）
元素の高度利用を基軸に新材料を効率的に探索するため、計算科学/データ科学/高スループット評価/非平衡プロセス/プロセス・インフォマティクスに直結させたその場計測などを含む材料創製手法を開発し、新機能の発見や、信頼性・耐久性の飛躍的な向上を実証することにより、元素高度利用の科学と新機能材料創製の開発基盤構築を目指す取り組みがなされている。
- ・JST 未来社会創造事業 / マテリアル探索空間拡張プラットフォームの構築（2021年～）
材料探索空間を圧倒的に拡張し、研究開発の効率を上げることを目標として、①ハイスループット自律探索システム、②データ駆動/仮説駆動ハイブリッド型研究、③ナレッジシェアリングの3つの柱でマテリアル探索空間拡張プラットフォームの構築を進めている。電池材料を題材とし材料探索スループットの1,000倍向上をPoC（Proof of Concept）に掲げている。
- ・JST 未来社会創造事業 / ファンデルワールス複合原子層の物性創発におけるマテリアルインフォマティクス活用と指導原理導出（2021年～）
ファンデルワールス複合原子層を題材として、マテリアルズ・インフォマティクス、プロセス・インフォマティクスを活用する。物性と構造の相関を明らかにして、社会実装に向けた指導原理導出を目指している。
- ・科研費 学術変革領域（A） / デジタル化による高度精密有機合成の新展開（2021年～）
有機合成の多様性に対応した独自のデジタル化プラットフォーム構築をめざす。反応条件最適化、合成経路探索、高次複雑系分子設計の3つの自動化システムを開発し、革新的な基礎反応の発掘や開発効率の超加速化（10倍以上）を実証することを目指している。また、バッチ反応からフロー反応への変換法開発と自律的な条件最適化ユニットを組み込んだ自動合成システムの構築にも取り組んでいる。
- ・NEDO マテリアル革新技術先導研究プログラム（2021年度）
先導研究によって15年から20年以上先の新産業創出に結びつく革新的なマテリアル技術のシーズを育成し、将来の国家プロジェクトなどにつなげることを目的としたマテリアル革新技術先導研究プログラムにおいて、データ駆動型物質・材料開発に関連する研究開発テーマとして「SiCバルク成長技術の革新に向けたプロセス・インフォマティクス技術の研究開発」「水分解水素製造用光触媒結晶のマテリアルDX研究開発」「データ駆動科学によるスマートスケラブルケミストリーの確立」「ファインセラミックスのプロセス・インフォマティクス基盤構築」が採択された。
- ・経産省 / 計測分析装置の計測分析データ共通フォーマット及び共通位置合わせ技術に関するJIS開発

(2020～22年)

NEDO「省エネ製品開発の加速化に向けた複合計測分析システム研究開発」(2018～19年)での検討を踏まえて、国内メーカーが強みを持つ電子顕微鏡や質量分析装置、X線分析装置などの各種計測分析機器のデータを統合させるプラットフォームを構築化(JIS化)し、AI等を活用した高度な解析を可能とする複合計測分析システムの開発を行う。

- ・JST-ERATO / 前田化学反応創成知能プロジェクト (2019年～)

有機合成化学における量子化学計算に基づいた反応経路探索を行い、情報科学的手法と組合せて、化学反応における原子の動きの全貌を予測し、有用な未知反応を次々に提案する「化学反応創成知能」を創出する。ロボットによる自動合成・データベース化にも取り組んでいる。

- ・SIP第2期 / 統合型材料開発システムによるマテリアル革命 (2018年～)

第1期SIP「革新的構造材料」後継プロジェクトとして発足し、マテリアルズインテグレーション(MI)の技術基盤を生かし、欲しい性能から材料・プロセスをデザインする逆問題MIに対応した統合型材料開発システムの開発を目指す。

- ・次期SIP(第3期)のフィージビリティスタディー (2022年)

大量に使用・廃棄されるプラスチック等素材を対象としたサーキュラーエコノミーシステムの構築、マテリアルユニコーンの継続的創出を目指すマテリアル・プロセス・イノベーション基盤技術構築が次期SIP課題候補として選定されており、それぞれにおいてデータ駆動型材料開発の推進が計画されている。

- ・JST-CREST / 実験と理論・計算・データ科学を融合した材料開発の革新 (2017年～)

これまで実施されてきた物質・材料開発の基本となる実験科学と、理論、計算、データ科学とを融合させることにより、革新的材料開発へとつながる手法の構築を目指す。

【海外】

- ・Materials Research Data Alliance (MaRDA) [米国]

2019年11月にスタートし、オープン・アクセス可能・相互運用可能な材料データを 実現することを目的として設立された。

- ・Center for Hierarchical Materials Design (CHiMaD) [米国]

CHiMaDは2014年からノースウエスタン大学を中心とする研究チームの5年プログラムのファンディングが開始し、NISTのMGI研究チームとの密接なコラボレーションのもとシカゴ地区でのデータ活用型材料研究を牽引してきたが、2019年にさらに5年間にわたるファンディングの継続が決定された。

- ・High Throughput Experimental Material Collaboratory (HTE-MC) [米国]

MGIの後継策の一つとして始まったHTE-MCは、高品質な実験データを迅速に大量取得してデータベースに登録し、材料設計を大幅に改善するための材料合成・特性化・データ管理サービスの統合型ネットワークである。

- ・Materials Innovation Platforms (MIP) [米国]

材料研究の進歩を加速するためにNSFがファンディングする拠点形成のプログラムであり、材料合成/加工 - 材料特性評価 - 理論/モデリング/シミュレーションを反復する“closed-loop”型研究を行うものである。第1期(2015年)では無機結晶(バルク、薄膜)の開発、第2期(2019年)では材料研究と生物科学の融合に焦点が当てられた。

(5) 科学技術的課題

日本では、実験データのデータ蓄積基盤が整っているとは言えない。それに比べ、米国では、NISTがHigh-Throughput Experimental Materials Collaboratory (HTE-MC)を立ち上げ、実験データの蓄積基盤を整えている。National Renewable Energy Laboratory (NREL)はPV関連材料などの実験デー

タの公開基盤を整備している。これらと、2011年よりMGIで進めてきた計算データ蓄積基盤（Material ProjectやAFLOWなど）を合わせることで、実験データと計算データの両方のデータ蓄積基盤が整うことになる。

日本においても、ハイスループット実験技術・コンビナトリアル実験技術・ロボット自律実験技術等を活用して実験データを効率的に蓄積する仕組みを整備することが、今後のデータ駆動型物質・材料開発において重要な因子になってくる。マテリアル先端リサーチインフラでのデータ蓄積をスムーズに立ち上げることが日本の競争力にとって重要である。

また、新奇物質・材料の合成には、合成過程のその場観察を通じた中間状態（反応中間体）を含めた理解を深めることが必要である。計測インフォマティクスなどを活用し、反応素過程の理解を深め、それに基づく高精度なシミュレーションを実現することも重要な課題である。大型共用研究施設SPring-8、J-PARC等でもデータ基盤の整備が始まっており、その活用が期待される。

(6) その他の課題

本研究開発領域は、複合領域の横断研究であり、物質・材料科学を中心として、機械学習・ロボット工学・制御理論の知識・技術の統合が必要となる。特に人材育成の面では、細分化された専門を超えて、横断的な関心を持ち、協業することが重要である。論文発表が主要な成果となる研究者のみならず、技術者、技能者の果たす役割は極めて重要であり、その処遇・キャリアパスは大きな課題である。

また、機械学習・AI技術は、材料開発のみならず、幅広い分野で研究開発が積極的に進められている。国内で複数あるプロジェクト間において、分野を超えた横断的な情報交換がなされ協働できることが望まれる。

自律実験システムにおいては、合成から計測・評価までを一貫した容器や搬送機構でシームレスにつなぐことが有効である。そのためには、研究開発の早い段階から、ロボティクスメーカーや計測機器メーカーと綿密に連携していくことが大切である。あわせて、システムの観点からIT/ソフト業界との連携強化も必要である。マテリアル分野はベンチャーが育ちにくいのが、データ駆動型マテリアル開発技術を活用することでベンチャー企業の創出が期待できる。

(7) 国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	↗	<ul style="list-style-type: none"> MI技術に関する論文数は増加傾向。 プロセス・インフォマティクス関連の論文も増加。 自律実験に関する研究も活発になってきている。
	応用研究・開発	○	↗	<ul style="list-style-type: none"> MI関連の産業界からのプレスリリースが大幅増加。 大企業だけでなく、ベンチャー企業（MI-6、Creative AI Roboticsなど）も健闘。 計測装置の規格標準化の動きもある。
米国	基礎研究	◎	↗	<ul style="list-style-type: none"> MI技術に関する論文数は増加傾向。 トランプ政権に比べ、バイデン政権ではMI関連技術への大型予算投入の可能性大。 自律実験に関する研究も増えてきている。
	応用研究・開発	◎	↗	<ul style="list-style-type: none"> 巨大IT企業（Google, Facebook, IBMなど）もMI分野に本格参入。 製薬関連では、大手企業（Eli-Lily）も自動開発プロセスを導入。自動開発プロセス自体の事業化の動きもある。（SRI international、KEBOTX社）

2.5

俯瞰区分と研究開発領域
物質と機能の設計・制御

欧州	基礎研究	◎	↗	<ul style="list-style-type: none"> MI技術に関わる論文数は増加傾向。 NOMADは追加予算で継続中。 自律実験については、英国を中心に大型プロジェクトが進行中。
	応用研究・開発	○	↗	<ul style="list-style-type: none"> 自動合成装置がChemspeed社より製品化。この商品を使った研究論文もある。 計測装置をはじめとした規格標準化・規制化については、欧州が主導するケースが目立つ。
中国	基礎研究	○	↗	<ul style="list-style-type: none"> MI技術に関わる論文数は増加傾向。 プロセスに関しても、有機化合物のフロー合成、MOFの開発などがでてきている。
	応用研究・開発	○	↗	<ul style="list-style-type: none"> MI関連特許多数
韓国	基礎研究	○	↗	<ul style="list-style-type: none"> MI技術に関わる論文数は増加傾向
	応用研究・開発	○	↗	<ul style="list-style-type: none"> MI関連特許多数

(註1) フェーズ

基礎研究：大学・国研などでの基礎研究の範囲

応用研究・開発：技術開発（プロトタイプの開発含む）の範囲

(註2) 現状 ※日本の現状を基準にした評価ではなく、CRDSの調査・見解による評価

◎：特に顕著な活動・成果が見えている

○：顕著な活動・成果が見えている

△：顕著な活動・成果が見えていない

×：特筆すべき活動・成果が見えていない

(註3) トレンド ※ここ1～2年の研究開発水準の変化

↗：上昇傾向、→：現状維持、↘：下降傾向

関連する他の研究開発領域

- AI・データ駆動型問題解決（システム・情報分野 2.1.6）
- AI創薬（ライフ・臨床医学分野 2.1.3）

参考・引用文献

- 国立研究開発法人科学技術振興機構研究開発戦略センター「データ科学との連携・融合による新世代物質・材料設計研究の促進（マテリアルズ・インフォマティクス）」<https://www.jst.go.jp/crds/report/CRDS-FY2013-SP-01.html>, (2022年12月22日アクセス)。
- 国立研究開発法人科学技術振興機構研究開発戦略センター「材料創製技術を革新するプロセス科学基盤～プロセス・インフォマティクス～」<https://www.jst.go.jp/crds/report/CRDS-FY2021-SP-01.html>, (2022年12月22日アクセス)。
- 文部科学省マテリアル先端リサーチインフラセンター運営室「文部科学省マテリアル先端リサーチインフラ」<https://nanonet.mext.go.jp/>, (2022年12月22日アクセス)。
- NSTC Subcommittee on the Materials Genome Initiative, “The 2021 Materials Genome Initiative Strategic Plan,” Materials Genome Initiative, <https://www.mgi.gov/sites/default/files/documents/MGI-2021-Strategic-Plan.pdf>, (2022年12月22日アクセス)。
- Materials Research Data Alliance (MaRDA), <https://www.marda-alliance.org/>, (2022年12月22日アクセス)。
- University of Chicago, “Globus,” <https://www.globus.org/>, (2022年12月22日アクセス)。
- AiiDA team, “AiiDA v2.0.0 released,” AiiDA, <https://www.aiida.net/news/aiida-v2-0-0-released/>, (2022年12月22日アクセス)。

- 8) European Materials Modelling Council (EMMC), “The EMMC Roadmap 2018 for Materials Modelling and Informatics,” https://emmc.eu/wp-content/uploads/2020/03/EMMC_Roadmap2018.pdf, (2022年12月22日アクセス) .
- 9) Rickard Armiento, “OPTIMADE v1.0.0,” GitHub, <https://github.com/Materials-Consortia/OPTIMADE/releases/tag/v1.0.0>, (2022年12月22日アクセス) .
- 10) Ryota Shimizu, et al., “Autonomous materials synthesis by machine learning and robotics,” *APL Materials* 8, no. 11 (2020) : 111110., <https://doi.org/10.1063/5.0020370>.
- 11) Sebastian Steiner, et al., “Organic synthesis in a modular robotic system driven by a chemical programming language,” *Science* 363, no. 6423 (2019) : eaav2211., <https://doi.org/10.1126/science.aav2211>.
- 12) Benjamin Burger, et al., “A mobile robotic chemist,” *Nature* 583, no. 7815 (2020) : 237-241., <https://doi.org/10.1038/s41586-020-2442-2>.
- 13) B. P. MacLeod, et al., “Self-driving laboratory for accelerated discovery of thin-film materials,” *Science Advances* 6, no. 20 (2020) : eaaz8867., <https://doi.org/10.1126/sciadv.aaz8867>.
- 14) Cynthia Rudin, “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,” *Nature Machine Intelligence* 1, no. 5 (2019) : 206-215., <https://doi.org/10.1038/s42256-019-0048-x>.
- 15) José Jiménez-Luna, Francesca Grisoni and Gisbert Schneider, “Drug discovery with explainable artificial intelligence,” *Nature Machine Intelligence* 2 (2020) : 573-584., <https://doi.org/10.1038/s42256-020-00236-4>.
- 16) Jinchao Feng, et al., “Explainable and trustworthy artificial intelligence for correctable modeling in chemical sciences,” *Science Advances* 6, no. 42 (2020) : eabc3204., <https://doi.org/10.1126/sciadv.abc3204>.
- 17) 国立研究開発法人科学技術振興機構未来社会創造事業「マテリアル探索空間拡張プラットフォーム (MEEP)」 <https://meep.nagato-u-tokyo.jp/>, (2022年12月22日アクセス) .