

## MI<sup>2</sup>I 拠点における物質科学データの整備と普及

伊藤 聡：物質・材料研究機構

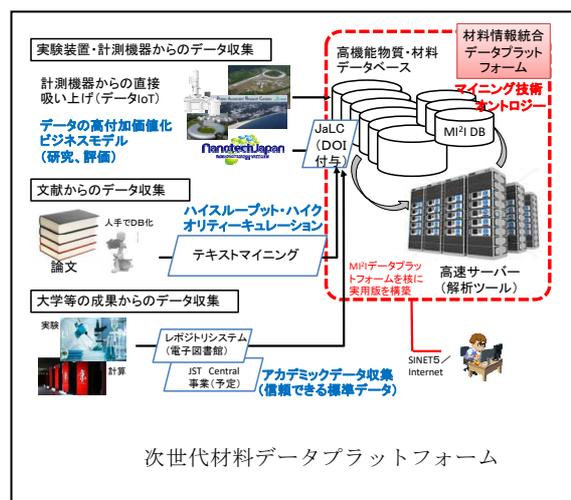
2015年7月より、JST イノベーションハブ構築支援事業の一つとして、物質・材料研究機構（NIMS）をハブ拠点に「情報統合型物質・材料開発イニシアティブ（MI<sup>2</sup>I）」が活動を進めています。MI<sup>2</sup>I は物性研究者やデータ科学研究者が参加するマテリアルズインフォマティクス（MI）の研究推進および体制構築プロジェクトです。ここでは、データ科学の手法を物性研究に生かす取り組みを進めています。具体的には磁石材料、蓄電池材料、伝熱制御材料に適用して、新しい蓄電池向け固体電解質探索や化合物半導体の熱伝導特性の最適化等を行っています。

マテリアルズインフォマティクスを進める上で、重要な要素の一つに材料データがあります。NIMS では 20 年近く、大規模な材料データベース MatNavi を整備・運用してきました。MatNavi はもともと人が利用することを前提とした高品質材料データベースとして設計・開発されています。しかしながらマテリアルズインフォマティクスの研究には、ソフトウェアからデータを呼び出せることが必要です。そのための API（Application Program Interface）を開発しています。また、データの蓄積に関しても、従来とは異なる手法を検討しています。

これまではデータの品質を重視し、特定のプロセスに依存しない物質データを中心に人手によってデータ収集を行ってきました。しかし、昨今のデータ科学に対応するにはより大量のデータを収集することが必要であり、かつ、プロセス情報も収集することが求められています。そのため、新しい ICT 技術を活用したデータ収集法の研究開発を行っています。材料研究を行う上で様々な実験データがとられていますが、現状ではその多くは研究者個人が抱え込んでおり、その大半は公表されることなく死蔵されています。そこで最新の IoT 技術を活用して、計測機器から直接、計測データをサーバーに収集し、キュレーションする仕組みの開発を進めています。キュレーションの段階でデータ科学に基づくスペクトル解析などの付加価値をつけることで、データ提供へのインセンティブになるように設計を進めています。

論文等で公表されたもののデータの質はある程度保証されています。しかしながら大量に出版される論文を人手でキュレーションすることは現実的ではなくなっています。これに対してはテキストデータマイニング（TDM）技術を利用することでキュレーションの効率化を進めています。文献から物質名を抽出する、図表からデータを抽出するといった物性分野特有の課題もあり、新しい TDM 技術の研究開発が必要となっています。

昨今、機関リポジトリがその機関の研究成果を収集しています。そこには貴重な材料データも収集されています。大型研究施設である放射光施設や NMR 施設などとも連携して、利用可能な材料データを収集・整備し、日本の科学技術基盤として長期運用していく予定です。



### 関連 web

<http://www.nims.go.jp/MII-I/index.html>