

前回（第7回）の議論における主な内容

（1）データ駆動型研究開発の促進（マテリアルDXプラットフォームの推進）

（データ駆動型研究開発推進）

- 多量な共通基盤（オープン）データが必要。特に初期の探索フェーズにおいて共通基盤データは必ずしも高品質である必要はなく、粗いデータでもよい。
- フェーズ毎に最適なデータの精度は異なる。例えば社会実装（産業化）のフェーズにおいては材料品質保証等の観点から従来型の実験手法による精度の高い検証が必要。
- 階層性・転移性を考慮したデータポートフォリオの構築が重要。例えば基盤データで大量なデータを作成し、限られたデータであるエッジデータとのブリッジは機械学習で行う。基盤とエッジデータを定める戦略が必要。
- データ基盤にはセキュリティや利用されたデータの追跡技術などが必要。
- 共通基盤データが集まりにくい要因として、材料空間は非常に広大で階層的であり共通基盤データをコミュニティで共創することが難しいことや、共有することにインセンティブが無いこと、材料研究者にとっては最先端（エッジ）材料・データが重要であること、そもそも材料分野はデータ取得が高コストであることが挙げられる。
- データを大量に創出するための自動実験が重要。一方で、自動実験設備を導入しても、低稼働率であったり、自動化システムがうまくいっていないケースがあり、設備を共有化するなど対策が必要。
- 自動実験を継続的にするためには、データ生産、価値創出、収益回収、再投資といったエコシステムの構築が必要。
- 研究リーダーがデータ研究駆動型研究に関するビジョンと具体的なシナリオを明示することが重要。

(情報系研究者の参画)

- 最先端の材料を生み出していくためには現状のデータ科学だけでは困難なため、計算データと実験データの両方が必要であり、材料研究者とデータ研究者の連携が必要。
- データ研究者の参画を促すには、開発したアルゴリズムを実践することができるオープンな共通基盤データや優れたユーザビリティ（一括ダウンロード可能やライブラリ機能など）をもったアプリなどの研究環境が必要。
（例：Material Project）
- AIやDXなどのルーツをもったデジタルネイティブなリーダーがいると求心力が生まれ、データサイエンティスト等情報系研究者が集まりやすい。
- データ研究者には大規模言語モデルや生成AIやバイオ、工学系など多様な選択肢があり、敢えてマテリアル分野を選択してもらうためには魅力やキャリアパスが必要。

(その他)

- データ駆動型研究開発が重要である一方でデータを必要としない研究者（データ科学を使わずとも最先端材料を生み出せる従来型研究）を無理に巻き込まないことが重要。得意分野でその力を発揮すべき。
- 産学官が蓄積してきたクローズドデータを集積するためには、データ数の増加により予測モデルの精度が向上することや、精度向上に必要なデータ数を定量的に示すことで、データ所有者の理解・協力を得ることが必要。

我が国の強み弱み

- ARIMは20年続いており、事業の大事さが認識されているということ。我が国の強みとして位置付けるべき。
- 我が国では、いわゆる「拠点・大型研究施設・大型国プロ参画機関」以外の一般的な材料科学研究室においても、比較的良質な研究装置を適切に維持・運用している（いた）。材料科学にとって大切な「ボトムアップ型研究の広がり」や「多様なシーズの創出」の点で、日本の強みと言える。
- 良質な研究装置を適切に維持・運用できる状況は、世界的装置メーカーとコンタクトがとりやすい国内環境にあることも一因。
- 最新鋭の機器は海外の研究機関・研究室が購入する傾向がある。我が国の一般的な材料科学研究室の研究費が、研究装置の高価格化にまったく追いついていない状況を反映しているのではないか。
- 材料科学にとっては、少数の大規模な拠点化とともに、比較的多数の中規模の拠点化も重要である。