

物質研究・材料開発分野のオープンサイエンス

～研究者・産業界がともにデータをオープンにできるインセンティブ・メカニズムの構築を～

1. 物質研究・材料開発はセレンディピティーのその先へ

松明を焚くことで明かりを得た人類は、その後、電気を制御する技術を知り、タングステンをフィラメントにすることで白熱電球を生み出し、蛍光体という材料を作り出したことで蛍光灯を手にし、暗闇を昼間と同じように利用することができるようになった。さらには半導体材料の開発によって各色のLEDが生み出され、長寿命・低消費電力の明かりが町中を明るくしている。人工衛星から撮られた地球の夜を見ると、この明かりの分布こそが人類が成し遂ってきた文明の広がりのように思われる。

このような新しい物質や材料は偶然に見つかることが多い。いわゆるセレンディピティーである。ひとたびそれが見つかれば、なぜ、それがそういう性質を持っているのか、それを研究する学問として物質科学が存在する。これは19世紀後半から急激に進み、いまでは結晶構造と元素の種類を決めれば、その物質は電気を通すか、通すならどのくらいの電気伝導度があるのかなどは定量的に予測できる。これを支えているのはそうした学問の進歩であり、コンピュータ能力の飛躍的向上である。こうしたアプローチを演繹的手法というが、それがあまりにも強力なため、そこで見落とされがちなことがあった。それは特性・機能から物質・材料を探索・設計するという演繹的手法の逆方向の立場である。セレンディピティー、それは研究者・技術者の勘と経験の素晴らしさを示すものではあるが、研究開発をそれで済ませて良いのだろうか？

2. 物質研究・材料開発でのオープンサイエンスの可能性

セレンディピティーは、はたから見ているからそう見えるのであって、本人にとっては十分な準備があった上での閃きであろう。十分な準備、それは過去の知見でありデータである。閃く人は多くの情報・データを持ち、そこに常人から見れば予想外のつながりを見つけることができるよう人である。これは常人には難しいことである。しかし、昨今のICT技術の進歩は埋もれていたデータを顕在化し、かつ到底人手では扱いきれないような膨大な量のデータを扱うことができるようになってきた。そこにはほとんど無価値のデータもあるだろうが（ごみはいくら集めても、ごみ）、しかし、量が違えば何かが変わることもある（都市鉱山の例もある）。こういう特性を持っている材料にはどんなものがあるかということを、膨大なデータを機械学習などの情報処理技術を使うことによって、従来の物質科学とは異なるアプローチで、目的に合致する最適な材料を見出すことが可能になりつつある。この分野は『マテリア

ルズ・インフォマティクス』と呼ばれる。一般には2011年の米国オバマ政権が打ち出したMaterials Genome Initiative (MGI)がその発端と言われているが、実はかならずしもそうではなく、とくに材料情報を集約し、データベース化して物質科学の研究者・技術者に広く提供することは日本が世界に先行して行ってきた。

マテリアルズ・インフォマティクスは材料開発の効率化・最適化という工学的な、すなわち産業上の価値が高いが、同時に物質科学に対する新しい視点を与えるという科学的な意義も大きい。これを推し進めるにはなによりも情報量が豊富で信頼性の高い材料データベースが必要である。これはオープンイノベーションとオープンサイエンスを支える基盤である。それ故、現在、国内外で材料データの収集が進められている。

3. オープンサイエンスを阻むもの

このような時流から考えると、材料データはどんどん集まり、マテリアルズ・インフォマティクスは学問として深まり、手段としては産学に普及していく、となるかというと、話はそう簡単ではない。オープン化を阻む要因があるのである。その一つは材料データの特異性である。材料情報には2種類あり、一つはハンドブックに掲載されているような物性データ、もう一つは作り方・プロセスデータである。前者はすでに人類共有の知見として広く収集・活用されているが、後者は数値化・定量化できないはずはないが、これまでむしろ意図的にそれを行わずに「匠の技」として秘匿されてきた感がある。例えば、ほとんど鉄元素から構成されているにもかかわらず、歴史上の名刀は刀鍛冶の技がわからないために現代技術をもってしても再現は困難である。そしてこのことは研究上の、あるいは産業上の競争力の源泉であり、差別化技術でもある。

他方でデータをオープンにする側のメリットが少ないことも問題である。ギブアンドテイクの精神でみんながデータを提供し、自分が持っていたデータとは違うデータを利用できることが、目指す世界であるが、残念なことにデータを提供した人が“出し損”になるのが現状である。さらには研究者の評価は原著論文によって行われているので、いくら優れたデータを提供しても研究としての評価がなされないという現実がある。産業界は当然、自社データは自社のために使うものであって、すでに使わなくなったデータであっても、別の用途で使われて利益を生み出す可能性があるとすれば、自社データを提供することは企業判断としてはありえない。

4. 競争と協奏の好循環に向けて

マテリアルズ・インフォマティクスは情報とモノとが密接に絡み、そこにノウハウや匠の技が関係している。それを担う人々は学問上も産業上も競争の場のプレーヤーである。そうした人々がデータを提供するほどにメリットを享受できることが一番のインセンティブであろう。たとえば、最近ではデータそのものを対象とした新しいタイプの学術誌（データジャーナル）が発刊されており、そこにデータを出すことで論文成果（研究成果）として評価される。自

社が提供したデータで他社が大きな成果を得、その結果、自社の競争力が弱まることは望ましくないので、こうした事態を回避する新しい制度も必要であろう。競争を協奏的に行うためにはその基盤を国が整備していく必要がある。その一方で、研究者・技術者はギブアンドテイクの気概を持って取り組んでいくことがオープンイノベーション・オープンサイエンスにおいて必要ではないだろうか？

※物質・材料研究機構のオープンサイエンスへの取り組みとして、物質・材料データベースの構築が進んでいます。

NIMS物質・材料データベース Mat Navi はこちら <http://mits.nims.go.jp/>

伊藤 聰（いとう さとし）

科学技術振興機構 情報統合型物質・材料開発イニシアティブ（MI²I）プログラムマネージャ
電機メーカー研究開発部門を経て、2011年理化学研究所、2015年から現職。
(公財)計算科学振興財団技術顧問等を兼任。専門分野は計算材料科学。
データ駆動型科学技術の推進と協奏的研究環境整備に取り組んでいる。

