

論文内グラフデータを収集・共有できる WEB システムの開発

A web system for collecting and sharing plot data on literature

理研¹, 東大院新領域², 物材機構³ ○熊谷 将也¹, 桂 ゆかり^{2,3}, 小谷 拓史^{2,3}, 大内 秀恭²,
郡司 咲子³, 安藤 有希³, 今井 庸二^{1,3}, 木村 薫², 津田 宏治^{1,2,3}

RIKEN¹, Univ. of Tokyo², NIMS³, ○Masaya Kumagai¹, Yukari Katsura^{2,3}, Takushi Kodani^{2,3},
Hideyasu Ouchi², Sakiko Gunji³, Yuki Ando³, Yoji Imai^{1,3}, Kaoru Kimura², Koji Tsuda^{1,2,3}

E-mail: masaya.kumagai@riken.jp

【緒言】従来の新規材料探索・設計は、過去の経験や知識などに基づいた概念駆動型アプローチが広く行われてきた。ところが最近では、材料応用の多様化、設計の複雑化が進み、これらの課題を解決する新たなアプローチへの期待が高まっている。中でも、膨大なデータから新たな規則性や知識を見出すデータ駆動型アプローチである Materials Informatics (MI)が急速な発展を遂げている。

現存する MI は、第一原理計算データを利用したものが多くを占めている。膨大なデータを必要とする MI にとって、第一原理計算は自動化によるデータ収集が容易であるため都合が良い。また、近年の計算技術・機器の発展に伴い、より短時間に膨大なデータが生成できるようになったことも理由の一つである。そして、その膨大なデータがインターネットを通して共有されたことで MI への関心が急速に高まり、新規材料探索・設計の促進に大きく貢献し始めている。ところが、理想的に設計、計算された材料の物性値は、実験で作製された材料の物性値としばしば異なった値を示す。第一原理計算では、実際の実験で生じる多くの相互作用を考慮しきれないためである。それらの相互作用を考慮するため、実験データを利用した MI の例 [1]も報告され始めているが、データそのものも含めて未だ極めて少ない。

そこで本研究では、実験データを利用した MI の促進のため、過去に出版された論文に掲載されている実験データを容易に抽出・収集、共有できるシステムの構築に取り組んだ。

【方法】実験データの共有に焦点を当て、WEB アプリケーションとしてシステムを構築した。データ抽出には、既存のデータ抽出ソフトウェア WebPlotDigitizer[2]をシステム内に導入した。データの収集対象には熱電変換材料を選択し、実験データの網羅的な収集を行った。

【結果】論文中のグラフから実験データを抽出・収集し、共有が可能な WEB アプリケーション”Starry data 2”の構築に成功した。実験データを MI で利用する際、実験内容などを考慮して前処理をすることが想定される。そこで、システム全体を論文情報整理ツールとしても利用できるように設計し、データの出力形式も前処理が容易な JSON (JavaScript Object Notation) 形式を採用した。これまでに 1028 試料 766 組成の熱電変換材料に関する物性値 (電気伝導率、熱伝導率、ゼーベック係数 等) の実験データを取得した。システムの全容および実験データの詳細は当日報告する。

【参考文献】

- [1] M.W. Gaultois et al., APL Mater. 4, 053213 (2016).
- [2] A. Rohatgi, WebPlotDigitizer (2011).