

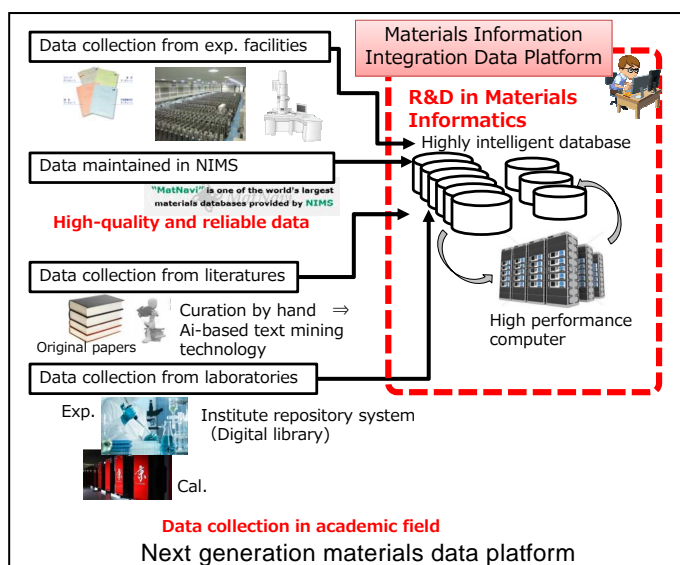
## A new materials data platform in MI<sup>2</sup>I

Satoshi ITOH : NIMS

The MI<sup>2</sup>I, which stands for Materials Research by Information Integration Initiative, has been started from July 1st 2015, as the JST project, and the NIMS, National Institute for Materials Science, is a hub organization of the MI<sup>2</sup>I. The MI<sup>2</sup>I is the first national project concerning the materials informatics in Japan, in which researchers in the fields of materials science and data science participate. In the MI<sup>2</sup>I, a new methodology on materials science by using data science has been studied. In particular, three subjects, which are magnetic materials, battery materials and thermal properties of materials, are considered, and for example, we are searching new solid electrolytes for lithium ion battery and optimizing thermal properties of compound semiconductors.

The materials data is one of important items to proceed the materials informatics. The NIMS has been constructing and operating a large-scale materials database called MatNavi over twenty year. The MatNavi was developed a high-quality materials database which was accessed via user-interface for researchers. However, it is needed to call data of the database from application software directly for materials informatics research, so that APIs (Application Program Interfaces), have been developed.

As a conventional scheme, we have been collecting materials data by hand with high quality, and nowadays, we have to accumulate a lot of materials data together with information on manufacturing processes for materials informatics study. For this sake, an accumulation scheme based on the present ICT technology has been developed. In this scheme, measurement data of instruments are uploaded to server semi-automatically. These data would be curated and analyzed



based on data science, which is designed for the merit of data providers.

Data in published papers are well-controlled, so that the papers are good resources of materials data. However, data curation by hand is not realistic because of a lot of publication of papers. We are going to accumulate materials data from literatures by using text-data mining technology, in which new text-data mining technology will be developed for material name extraction, data extraction from figures, and so on.

Several institute repository systems are operating and archives their research outputs. We will be accumulating materials data in cooperation with such organizations together with synchrotron radiation and/or NMR facilities, as an intelligence infrastructure in Japan.

---

### External links

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