INTERCONNECTING DATA REPOSITORIES: THE PLATFORM MATERIAL*DIGITAL* (BMBF)

Birgit Skrotzki, Thilo Muth and Pedro D. Portella

Bundesanstalt für Materialforschung und -prüfung (BAM)
D-12200 Berlin, pmd@bam.de, www.bam.de
www.materialdigital.de

Key Words: *Material data management, Standardization, Interoperability, Ontology, Workflow environments.*

Funded by the German Federal Ministry of Education and Research (BMBF), the Platform Material*Digital* aims at the development of a sustainable platform that brings together and supports interested parties from industry and academia in the digital transformation of Materials Science and Engineering [1]. This joint project between the Fraunhofer Institute for Mechanics of Materials (IWM), the Karlsruhe Institute for Technology (KIT, a member of the Helmholtz Association), the Leibniz Institute for Materials Engineering (IWT), the Max-Planck-Institut für Eisenforschung (MPIE) and the Bundesanstalt für Materialforschung und -prüfung (BAM) will develop initial approaches for the necessary complex data management, including contributions from all sectors working with materials data, such as companies, non-university research institutes and universities. By implementing this concept, a standardized exchange of data, knowledge and tools for the digitalization of materials that functions across domain boundaries is to be achieved.

In the second phase of this initiative different research projects start currently (https://www.materialdigital.de/projects/); each of these projects aims at the digital description of the complete cycle from the design to the production and characterization of a specific material – in a broad spectrum from high strength flat steels over inorganic glasses to Heusler alloys for magnetocaloric applications. In the subsequent phases of the initiative the materials data space will be progressively covered in similar projects with different funding sources.

Data-driven procedures in materials research will benefit enormously from the use of data from different sources. Their combination requires a standardization of data formats, machine-readable metadata, and the extracted information. In view of the abundance of different data sources and their broadness, a specific structuring of the data and associated metadata and information must be developed by the respective domain experts. In order to be able to take international advances in materials research and the internationally fragmented value chains in industry into account, a worldwide exchangeability and interconnectivity - interoperability - is necessary. The approach for ontology and standardization of the description of material data will be compatible with multinational initiatives such as the European Materials Modelling Council (EMMC) right from the very beginning.

The Material*Digital* platform will coordinate and advance the development of standardized software tools for the generation, storage, management, analysis, and distribution of material data. The objective of these activities is to provide both data and service providers and industrial users with a standardized, flexible, and easy-to-use toolbox for implementing digitalization projects. The development of these tools considers the wide range of material science questions, the multitude of available data sources and simulation and analysis methods, but also the dynamic development of the field. In order to enable their inclusion in industrial processes, the

interests of industrial users with regard to the confidentiality and security of their data and queries are already taken into account in the design of the tools.

In its initial version the Platform will provide two workflow environments called "pyiron" and "SimStack" in order to establish digital workflows - decentralized data or simulation concepts - by active agents within its software environment.

pyiron [2, 3] is a Python based framework which provides all the tools needed to interactively explore, implement, and run complex simulation protocols that require to combine different computer codes and to run thousands of separate calculations on high-performance computer clusters. pyiron allows to interactively implement and test simulation protocols and to upscale them for high-throughput simulations on large computer clusters.

SimStack [4] facilitates the efficient implementation, adoption and execution of complex and extensive simulation workflows and enables fast uptake of modelling techniques for advanced materials by industry. Software and workflow developers can incorporate their modules into SimStack, providing automatically generated graphical user interface for complex scientific code to the end-user.

With regard to the sustainability of the overall initiative, the Platform distinguishes between tasks that can be implemented in projects of stakeholders of different constellations and structural key and service tasks that, as enablers, should promote and accelerate the implementation of the projects in a standardized environment. On the basis of a holistic concept that includes both structuring elements and tools, the Platform develops a sustainable business model for the economic connectivity of the platform beyond the funding period.

REFERENCES

[1] https://www.bmbf.de/upload_filestore/pub/MaterialDigital.pdf

[2] Janßen, J.; Surendralal, S.; Lysogorskiy, Y.; Todorova, M.; Hickel, T.; Drautz, R.; Neugebauer, J.: pyiron: An integrated development environment for computational materials sciences. Computational Materials Science **163** (2019) 24 - 36 https://doi.org/10.1016/j.commatsci.2018.07.043

[3] https://www.mpie.de/3880353/pyiron

[4] http://www.simstack.eu/