



IUC04: Towards semi automated construction of Defect Phase Diagrams

IUC members: Khalil Rejiba¹, Prince Mathews², Niklas Siemer², Ulrich Kerzel³, Pavlina Kruzikova⁴, Ujjal Saikia⁴, Philipp Ost⁵, Sandra Korte-Kerzel¹, Tilmann Hickel⁴

DOI: 10.5281/zenodo.13220179

¹ RWTH Aachen University, Institute for Physical Metallurgy and Materials Physics

² Max-Planck-Institut für Nachhaltige Materialien GmbH

³ RWTH Aachen University, Fakultät für Georessourcen und Materialtechnik, Data Science and Artificial Intelligence in Materials and Geoscience

⁴ Bundesanstalt für Materialforschung und -prüfung (BAM)

⁵ Karlsruhe Institute of Technology (KIT), Scientific Computing Center (SCC)

Since defect phase diagrams are a novel concept in materials science, only a limited amount of thermo-chemo-structural data is currently available. They need to be attributed to certain defect types, defect characters and defect states. The main goal of IUC 04 is to ensure a model-driven (guided probing) collection of the relevant experimental and computational data to construct these diagrams and their post-processing according to newly established simulation protocols.

The collection of relevant data requires a framework where all required data can be stored. Furthermore, the data needs to be annotated by metadata to allow for searching of relevant data. We use OpenBIS as a central data management tool. It acts as a metadata store referencing the data stored in Coscine S3 storage. Finally, the workflow to construct defect phase diagrams is implemented with pyiron, an integrated development environment for computational materials science.

Currently, we work on two aspects of our general mission simultaneously. The entry of data into the system, and the workflow to construct defect phase diagrams from the data. The current status of data entering into the system, i.e. the aspects of documentation in electronic lab notebooks, the actual upload of data into the system, and the automated metadata extraction is discussed in the following.

Electronic Lab Notebook (ELN)

Custom ELN entries were designed in accordance with requirements collected from subject matter experts to capture metadata not present in the files generated by the scientific instruments. It is planned to introduce more custom ELN entries and expand the already existing metadata schemata as the need arises since they allow the definition of standard

operating procedures ensuring better reproducibility and easier onboarding of new researchers.

Data Upload into Cloud Storage

The laboratory information management system (LIMS) is extended with external cloud storage (Coscine S3). At the moment, two alternatives for uploading data to the cloud storage can be used. As a script-based approach is geared more towards advanced users, a graphical user interface was developed to enable the general user to upload data while linking it to the corresponding ELN entry. A plan to extend these capabilities to allow more automation is being explored. In parallel, we plan to adapt the script-based approach to work with the pyiron simulation framework to allow for an upload of calculations directly from pyiron including data and metadata.

Automatic Metadata Extraction

Metadata parsers for Scanning Electron Microscopy data and Electron Backscatter Diffraction data were developed. Consolidation of the metadata fields that correspond to the same physical parameter was performed in order to reduce vendor-specific nomenclature. Categorization of expected data types was conducted based on feedback from researchers. Open-source packages that handle scientific data were identified and examined. These packages were used to develop preliminary parsers for Transmission Electron Microscopy and Optical Microscopy. The progressive introduction of parsers in the upload tools is planned as they become available.

Defect Phase Diagram Construction

Another aspect of our work is the construction of defect phase diagrams from the collected data. To this end, an automatic workflow for the construction of defect phase diagrams based on ab-initio simulations at 0 Kelvin has been developed using the pyiron framework. Currently, this workflow can be used for binary systems and has been demonstrated for the example defect system of $\Sigma 7$ (12-30) [0001] 21.78° Mg symmetric tilt grain boundary. For this system, experimental findings performed in CRC 1394 (PP02 in NFDI-MatWerk) suggested different structural motives with and without Ga impurities in the system. In addition, the workflow has been extended to construct defect phase diagrams at finite temperatures.

Acknowledgements

This publication was written by the NFDI consortium NFDI-MatWerk in the context of the work of the association German National Research Data Infrastructure (NFDI) e.V.. NFDI is financed by the Federal Republic of Germany and the 16 federal states and funded by the Federal Ministry of Education and Research (BMBF) – funding code M532701 / the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the National Research Data Infrastructure – NFDI 38/1 – project number 460247524.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of Collaborative Research Centre CRC 1394 - Structural and Chemical Atomic Complexity - From Defect Phase Diagrams to Material Properties (project number 409476157).