High-Throughput Materials Informatics Integrating Ab Initio, Machine Learning and CALPHAD Data

I.Roslyakova[1], C. Früh[1], C. Aras[1], Ö. Büyükuslu[1], M. to Baben[1]
[1]GTT-Technologies, Kaiserstraße 103, 52134 Herzogenrath, Germany

ir@gtt-technologies.de, cf@gtt-technologies.de

CALPHAD modelling provides a rigorous framework for predicting phase stability and thermodynamic properties across temperature, pressure, and composition spaces. However, the development of comprehensive thermodynamic databases remains resource-intensive and constrained by limited experimental data, restricting rapid exploration of large compositional spaces.

To address this challenge, we present a machine learning-enhanced, high-throughput approach that leverages the aiMP and aiOQ databases to bridge first-principles calculations (DFT) and CALPHAD-based materials informatics workflows [1].

The aiMP and aiOQ databases are ab initio-derived thermodynamic resources developed by GTT-Technologies using 0 K DFT data from the Materials Project (MP) [2] and the Open Quantum Materials Database (OQMD) [3]. The aiMP database includes stable phases for 136 854 compounds, while the aiOQ database contains data for 475 887 compounds, together enabling extensive coverage of composition space unavailable in traditional CALPHAD databases. The aiMP and aiOQ databases integrate machine learning predictions of temperature-dependent entropy and heat capacity, derived from large-scale datasets comprising thousands of compounds, thereby extending 0 K DFT formation energies to finite temperatures and enabling CALPHAD-compatible Gibbs energy descriptions [1].

These databases enable three key applications:

1. **Data Cleaning & Reliability Assessment:** Automated anomaly detection in curated and ab initio data, ensuring more robust predictions for CALPHAD modelling.

2. **High-Throughput Alloy Screening:** Enabling rapid evaluation of wide compositional ranges to identify promising material candidates for further computational and experimental investigation.

3. **Phase Diagram Estimation and Optimization:** Providing efficient starting points for phase diagram development through CALPHAD optimization workflows.

Together, aiMP and aiOQ enable scalable exploration of complex compositional spaces beyond the limits of traditional database development. To illustrate the practical impact, we demonstrate the high-throughput alloy screening by leveraging the aiMP/aiOQ databases within a CALPHAD-based materials informatics workflow to identify stable coating candidates for high-temperature oxidation resistance of SiC claddings.

References

[1] C. Frueh, C. Aras, Ö. Büyükuslu and M. to Baben, Calphad, 90 (2025) 102838.
[2] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K.A. Persson, APL Materials, 1 (2013) 011002.
[3] J.E. Saal, S. Kirklin, M. Aykol, B. Meredig and C. Wolverton, JOM, 65 (2013) 1501–1509.