ULtrahigh TEmperature Refractory Alloys (ULTERA) Database Infrastructure and Data Quality Assurance

*A. M. Krajewski, A. Debnath, S. Lin, M. Ahn, H. Sun, W. Reinhart, A. Beese1, Z.-K. Liu

ULTERA database, developed under the ARPA-E's ULTIMATE program, is aimed at collecting literature data on high entropy alloys (HEAs) to facilitate rapid discovery of new ones using forward and inverse design. Our primary focus is ultra-high-temperature applications, which include more efficient jet engines which could bring us closer to a sustainable future.

As of September 2022, ULTERA contains over 6,200 property-datapoints, corresponding to 2,485 unique HEAs, collected from 455 source DOIs. All data is available through a high-performance API, following FAIR principles, while statistics on it can be found on our phaseslab.com/ultera web page. The database architecture is designed to automatically integrate starting literature data in real time with methods such as experiments, generative modeling, predictive modeling, and validations.

Here, we showcase the database with particular emphasis on our methods developed to assure high quality of data and screen for abnormal entries, a step crucial to creating a dataset for data-driven alloy design.