

Supplementary Information:

Unsupervised Data Mining in nanoscale X-ray Spectro- Microscopic Study of NdFeB Magnet

Xiaoyue Duan¹, Feifei Yang², Erin Antono³, Wenge Yang⁴, Piero Pianetta⁵,
Stefano Ermon³, Apurva Mehta^{5,*}, Yijin Liu^{5,*}

¹School of computer, Wuhan University, Wuhan, Hubei 430072, China.

²Division of Biomaterials and Bioengineering, Department of Preventive and Restorative Dental Sciences, UCSF, San Francisco, CA 94143-0758, USA

³Department of Computer Science, Stanford University, Stanford, CA 94305-2205, USA.

⁴Center for High Pressure Science and Technology Advanced Research, Shanghai 201203, China.

⁵Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA.

*Correspondence and requests for materials should be addressed to Y.L. (liuyijin@slac.stanford.edu) and A.M. (mehta@slac.stanford.edu)

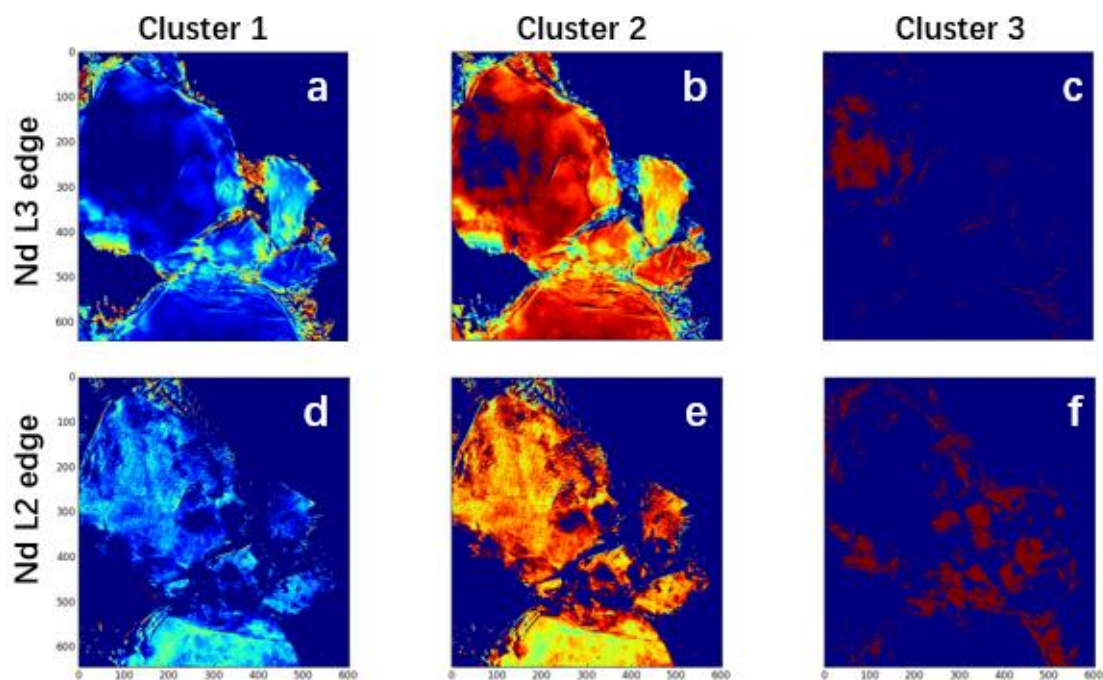


Figure S1. Cluster maps generated by applying PCA. The maps associated with the Nd L3 edge data are shown in panels a, b, and c. The maps associated with the Nd L2 edge data are shown in panels d, e, and f, respectively. Significant differences can be observed between the top row and the bottom row, indicating the PCA clustering failed to satisfy the consistency between the Nd L3 and L2 edges in the presented case study, which is a known physical constrain.

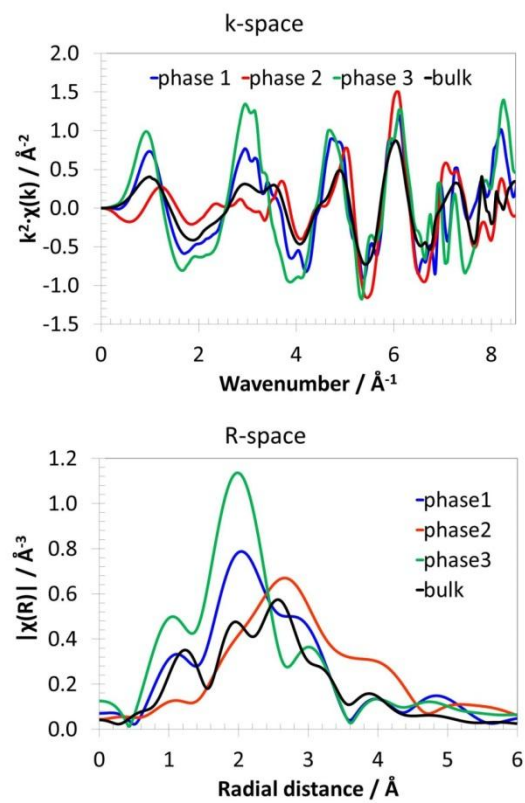


Figure S2. EXAFS analysis of the Nd L3 edge spectrums the three chemical phases as identified by the DBSCAN clustering algorithm and the bulk spectrum. Panel a: the k^2 weighted EXAFS signals in the k-space; panel b: R-space plot calculated.