Materials Informatics and Big Data:
Realization of 4th Paradigm of Science in Materials Science

Ankit Agrawal,* Alok Choudhary

Department of Electrical Engineering and Computer Science, McCormick School of Engineering and Applied Science, Northwestern University, USA

*Email: ankitag@eecs.northwestern.edu

Abstract: In this age of “big data”, large-scale experimental and simulation data is increasingly becoming available in all fields of science, and materials science is no exception to it. Our ability to collect and store this data has greatly surpassed our capability to analyze it, underscoring the emergence of the fourth paradigm of science, which is data-driven discovery. The need to use of advanced data science approaches in materials science is also recognized by the Materials Genome Initiative (MGI), further promoting the emerging field of materials informatics. In this talk, I would present some of our recent works employing state-of-the-art data analytics including deep learning for exploring processing-structure-property-performance (PSPP) linkages in materials, both in terms of forward models (e.g. predicting property for a given material) and inverse models (e.g. discovering materials that possess a desired property). Examples of forward models include predicting mechanical properties such as fatigue strength and microscale strain distribution, thermodynamic properties such as stability, and thermoelectric properties such as Seebeck coefficient. Examples of inverse models include discovery of stable compounds, indexing electron back-scatter diffraction (EBSD) patterns, and microstructure optimization of a magnetostrictive Fe-Ga alloy. I will also demonstrate some online web-tools we have developed that deploy machine learning models to predict materials properties. Such data-driven analytics can significantly accelerate prediction of material properties, which in turn can accelerate the optimization process and thus help realize the dream of rational materials design. The increasingly availability of materials databases along with groundbreaking advances in data science approaches offers lot of promise to successfully realize the goals of MGI, and aid in the discovery, design, and deployment of next-generation materials.

References:


Funding:

NIST Award No. 70NANB14H012, AFOSR Award No. FA9550-12-1-0458, DARPA Award No. N66001-15-C-4036, and Northwestern Data Science Initiative.