Chapter 7
Informatics–Based Approaches for Accelerated Discovery of Functional Materials

P. V. Balachandran
Los Alamos National Laboratory, USA

J. M. Rondinelli
Northwestern University, USA

ABSTRACT

This chapter is aimed at readers interested in the topic of informatics-based approaches for accelerated materials discovery, but who are unfamiliar with the nuances of the underlying principles and various types of powerful mathematical tools that are involved in formulating structure–property relationships. In an attempt to simplify the workflow of materials informatics, we decompose the paradigm into several core subtasks: hypothesis generation, database construction, data pre-processing, mathematical modeling, model validation, and finally hypothesis testing. We discuss each task and provide illustrative case studies, which apply these methods to various functional ceramic materials.

INTRODUCTION

Advancements in materials synthesis, processing, characterization, theory, and computation have significantly changed the way in which we generate, share, and analyze data in materials science. Modern day combinatorial experiments enable routine screening of large areas of composition space in a high-throughput fashion generating large volumes of data in an accelerated manner (Koinuma & Takeuchi, 2004; Potyrailo *et al.* 2011). In the arena of computational materials research, the last two decades have also seen tremendous progress in high-performance and high-throughput *ab initio* quantum mechanical-based calculations, enabled by ever increasing computational power at lower costs and improvements in parallel and optimization algorithms. The culmination of these efforts is a new virtual platform for the discovery, design, and development of functional materials and devices (Curtarolo *et al.* 2012; Jain *et al.* 2013; Saal *et al.* 2013). They are used in more than 10,000 publications annually, providing a bal-

ance between accuracy and efficiency, solving the *N*-electron Schrodinger equations primarily *via* the density-functional theory (DFT) formalism (Hohenberg & Kohn, 1964; Kohn & Sham, 1965). They have an outstanding track record of enabling theorists to predict properties of materials that are subsequently validated experimentally (Parr & Wang, 1995; Ceder, 1998; Rondinelli *et al.* 2008 & 2012; Mannhart & Scholm, 2010; Setyawan & Curtorolo, 2012; Olson, 2013).

Aided by the advancements in various levels of materials theory, including methods based on Green's functions (GW approach, Hedin, 1965) and dynamical mean field theory (DMFT, Georges *et al.* 1996; Kotliar *et al.* 2006) for excited state and correlated electron materials problems, improved exchange-correlation functionals, user-friendly computational codes [such as VASP (Kresse & Furthmuller, 1996; Kresse & Joubert, 1999) and Quantum ESPRESSO (Gianozzi *et al.* 2009) to name a few], and the growing infrastructure of massive supercomputing user-facilities and cloud computing, high-throughput and accurate *ab initio* calculations are generating volumes of data at an unprecedented level. For instance, digital databases such as "The Materials Project" (Jain *et al.* 2013), "Aflowlib.org" (Curtorolo *et al.* 2012) and "OQMD.org" (Saal *et al.* 2013) store *ab initio* calculations performed on more than 38,151, 17,200 and 284,000 entries, respectively, and counting.

Despite the encouraging progress, we have only scratched the surface when it comes to navigating the vast chemical space of inorganic compounds. There is a massive amount of compounds remaining to be explored beyond those reported in experimental structure databases. Multicomponent inorganic systems, which are the focus of this chapter, form the crux of modern materials science – inorganic compounds with more than two cationic sites appear to be the minimal requirement (Cheetham *et al.* 1994) for crystalline materials with functional electronic, magnetic, and optical properties. The flexibility of atomic structures allow fine tuning of these properties by altering, (i) interatomic distances (by using cations with carefully selected ionic radii), (ii) degree of covalency (by judicious choice of elements) and (iii) active frontier orbitals (by varying the oxidation states of the cations and anions contributing states to the low energy electronic structure in a controlled manner).

The attractive structural and functional features, which arise upon going from a binary to ternary to higher-order multi-element compound, however, introduces an enormous challenge in the rational search and discovery paradigm. To understand the complexity of the problem, consider the 60 elements highlighted in the periodic table (see Figure 1). Consider the case of a three-component (ternary) *ABC* materials system, where *A*, *B* and *C* are one of the 60 highlighted elements in Figure 1. Then, there are $C_3^{60} = 34,220$ possible combinations. In addition, Muller and Roy have documented more than 20 different major structure-types in their compendium on *The Major Ternary Structural Families* (Muller & Roy, 1974). This increases the number of possible candidates to 0.68×10^6 combinations. Furthermore, three-dimensional crystallography indicates there are 230 space groups in which the atoms can be periodically and uniquely arranged. Even if we intelligently refine our choice to (on an average) 5 space groups per candidate composition, then the number of calculations increases to 3.4×10^6 structures!

Now, consider the possibility of solid solutions ($BaTiO_3$-$SrTiO_3$, where the Sr and Ba atoms are distributed randomly in the unit cell lattice), defects ($La_2NiO_{4+\delta}$, where δ is the excess interstitial oxygen atoms in the lattice), and ordered cationic sublattices—the chemical space explodes. Even with the most powerful supercomputers, many compositions remain beyond the scope of *ab initio* calculations in part due to the large number of atoms in the unit cell. It is well known that the complexity of DFT calculations scale on the order of N^3, where *N* is the number of electrons (Parr & Wang, 1995). Informatics-based approaches have the potential to aid both high-throughput experimental and computational studies, by

Related Content

Study on the Ultrasonic-Assisted Vibration Tapping Using Automatic Tracing Frequency System
K. L. Kuoand C. C. Tsao (2014). *International Journal of Materials Forming and Machining Processes (pp. 73-87).*
www.irma-international.org/article/study-on-the-ultrasonic-assisted-vibration-tapping-using-automatic-tracing-frequency-system/106960

Pre-Test and Analysis of a Reinforced Concrete Slab Subjected to Blast From a Non-Confined Explosive
Fausto B. Mendonçaand Girum S. Urgessa (2018). *Energetic Materials Research, Applications, and New Technologies (pp. 272-287).*
www.irma-international.org/chapter/pre-test-and-analysis-of-a-reinforced-concrete-slab-subjected-to-blast-from-a-non-confined-explosive/195309

Optimization of Hot Extrusion Process Parameters Using Taguchi Based Grey Relation Analysis: An Experimental Approach
Sarojini Jajimoggala (2019). *International Journal of Materials Forming and Machining Processes (pp. 1-18).*
www.irma-international.org/article/optimization-of-hot-extrusion-process-parameters-using-taguchi-based-grey-relation-analysis/221322

Laser Metal Deposition Process
Rasheedat M. Mahamood (2016). *Advanced Manufacturing Techniques Using Laser Material Processing (pp. 46-59).*
www.irma-international.org/chapter/laser-metal-deposition-process/149837

Application of Teaching: Learning Based optimization to Surface Integrity Parameters in Milling
Nandkumar N. Bhopale, Nilesh Nikamand Raju S. Pawade (2015). *International Journal of Materials Forming and Machining Processes (pp. 1-16).*
www.irma-international.org/article/application-of-teaching/130695