

# Table of Contents

1. New tool in the box .....	3
2. Prediction of the atomic structure and thermoelectric performance for semiconducting $\text{Ge}_1\text{Sb}_6\text{Te}_{10}$ from DFT calculations .....	3
3. Generative deep learning as a tool for inverse design of high entropy refractory alloys .....	4
4. Generative models for inverse design of inorganic solid materials .....	5
7. Informatics is fueling new materials discovery .....	7
8. Composition genes in materials .....	7
9. Machine learning-guided design and development of metallic structural materials .....	8
10. Taking materials dynamics to new extremes using machine learning interatomic potentials .....	9
11. Boosting for concept design of casting aluminum alloys driven by combining computational thermodynamics and machine learning techniques .....	10
12. Big data-assisted digital twins for the smart design and manufacturing of advanced materials: from atoms to products .....	11
13. A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability .....	13
14. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments .....	13
15. Domain knowledge-guided interpretive machine learning: formula discovery for the oxidation behavior of ferritic-martensitic steels in supercritical water .....	15
16. Structure-property modeling scheme based on optimized microstructural information by two-point statistics and principal component analysis .....	16
17. Giant rashba splitting of confined Te chains in nanotubes: the size-, chirality-, and type-effects of nanotubes .....	17
18. Recent progress in the data-driven discovery of novel photovoltaic materials .....	18
19. Estimating the performance of a material in its service space via Bayesian active learning: a case study of the damping capacity of Mg alloys .....	19
20. High-cycle fatigue S-N curve prediction of steels based on a transfer learning-guided convolutional neural network .....	20
21. Accurate experimental band gap predictions with multifidelity correction learning .....	21
22. A thermodynamic database of the Ni-Mo-Re system .....	21
23. Additive manufacturing as a tool for high-throughput experimentation .....	22
24. CuB monolayer: a novel 2D anti-van't Hoff/Le Bel nanostructure with planar hyper-coordinate boron/copper and superconductivity .....	23

25. A mini review of machine learning in inorganic phosphors .....	24
26. Machine learning-accelerated first-principles predictions of the stability and mechanical properties of L1 <sub>2</sub> -strengthened cobalt-based superalloys .....	25
27. Process parameter optimization of metal additive manufacturing: a review and outlook .	26
28. A metadata schema for lattice thermal conductivity from first-principles calculations .....	27
29. New trends in additive manufacturing of high-entropy alloys and alloy design by machine learning: from single-phase to multiphase systems .....	28
30. High-entropy alloy catalysts: high-throughput and machine learning-driven design .....	29

## Editorial

### 1. New tool in the box

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Zhang, T.Y. New tool in the box. *J. Mater. Inf.* 2021, 1, 1. <http://dx.doi.org/10.20517/jmi.2021.01>

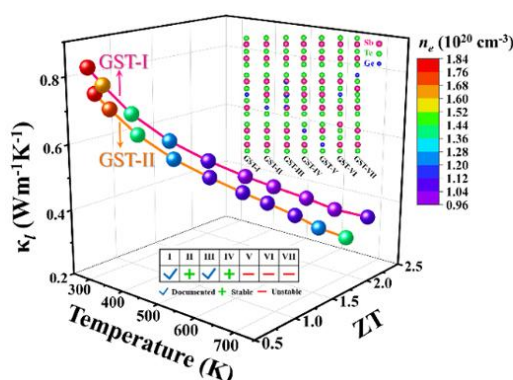
## Research Article

### 2. Prediction of the atomic structure and thermoelectric performance for semiconducting $\text{Ge}_1\text{Sb}_6\text{Te}_{10}$ from DFT calculations

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Gan, Y.; Zhou J.; Sun Z. Prediction of the atomic structure and thermoelectric performance for semiconducting  $\text{Ge}_1\text{Sb}_6\text{Te}_{10}$  from DFT calculations. *J. Mater. Inf.* 2021, 1, 2. <http://dx.doi.org/10.20517/jmi.2021.03>

#### Abstract



Pseudobinary alloys  $(\text{GeTe})_m(\text{Sb}_2\text{Te}_3)_n$  (GST), known as the most popular phase change materials for data-storage applications, also exhibit great potential as thermoelectric (TE) materials due to their intrinsically low lattice thermal conductivity ( $\kappa_l$ ) and high electrical conductivity. Among the GST compounds, the  $\text{Sb}_2\text{Te}_3$ -rich  $\text{Ge}_1\text{Sb}_6\text{Te}_{10}$  ( $m = 1$  and  $n = 3$ ) crystallizes into a complex trigonal structure with a 51-layer long period stacked along the  $c$ -axis, which may generate various possible atomic arrangements, thereby affecting the electronic and transport properties. Here, using *ab initio* calculations, we demonstrate that, besides the two experimentally known atomic sequences (GST-I and GST-III),  $\text{Ge}_1\text{Sb}_6\text{Te}_{10}$  has two novel stable stacking configurations (GST-II and GST-IV). GST-IV exhibits semi-metallic behavior, whereas GST-I and GST-II are semiconductors. Both semiconducting stackings have low  $\kappa_l$  of 0.86 and 0.78  $\text{W m}^{-1} \text{K}^{-1}$  at 300 K, owing to their small phonon group velocities and short phonon lifetimes. Moreover, they show a combination of high  $n$ -type Seebeck coefficient and electrical conductivity due to the steep slope of conduction band density of

states near bandgap, multiple conduction pocket electrons, and multiband conduction. The maximum ZT values of 2.23 and 1.91 are achieved in *n*-type stackings GST-I and GST-II at 710 K. Our work sheds light on the great potential of Ge<sub>1</sub>Sb<sub>6</sub>Te<sub>10</sub> with different atomic stackings for TE applications and will stimulate further experimental study. More importantly, from the perspective of materials informatics, this study provides significant insights that crystal systems with multilayered structures may open a viable route for creating new functional materials.

## Keywords

Ge<sub>1</sub>Sb<sub>6</sub>Te<sub>10</sub>, atomic stacking, structural stability, transport properties, thermoelectric performance

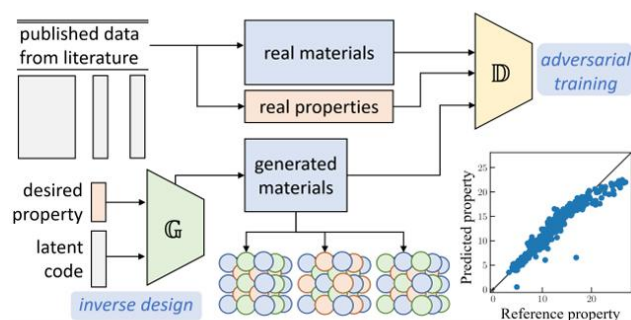
## Perspective

### 3. Generative deep learning as a tool for inverse design of high entropy refractory alloys

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Debnath, A.; Krajewski AM.; Sun H.; Lin S.; Ahn M.; Li W.; Priya S.; Singh J.; Shang S.; Beese AM.; Liu Z.K.; Reinhart WF. Generative deep learning as a tool for inverse design of high entropy refractory alloys. *J. Mater. Inf.* 2021, 1, 3. <http://dx.doi.org/10.20517/jmi.2021.05>

## Abstract



Generative deep learning is powering a wave of new innovations in materials design. This article discusses the basic operating principles of these methods and their advantages over rational design through the lens of a case study on refractory high-entropy alloys for ultra-high-temperature applications. We present our computational infrastructure and workflow for the inverse design of new alloys powered by these methods. Our preliminary results show that generative models can learn complex relationships to generate novelty on demand, making them a valuable tool for materials informatics.

## Keywords

High entropy alloys, databases, machine learning, inverse design

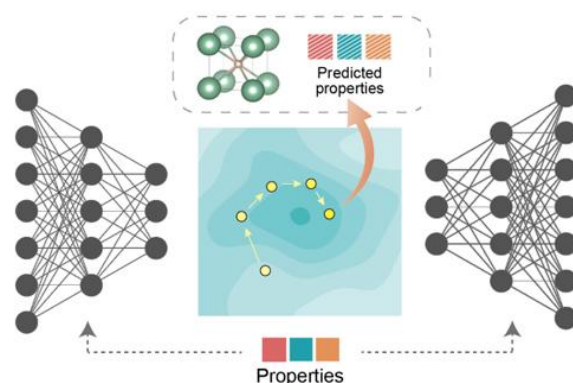
## Review

### 4. Generative models for inverse design of inorganic solid materials

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Chen, L.; Zhang W.; Nie Z.; Li S.; Pan F. Generative models for inverse design of inorganic solid materials. *J. Mater. Inf.* 2021, 1, 4. <http://dx.doi.org/10.20517/jmi.2021.07>

## Abstract



Overwhelming evidence has been accumulating that materials informatics can provide a novel solution for materials discovery. While the conventional approach to innovation relies mainly on experimentation, the generative models stemming from the field of machine learning can realize the long-held dream of inverse design, where properties are mapped to the chemical structures. In this review, we introduce the general aspects of inverse materials design and provide a brief overview of two generative models, variational autoencoder and generative adversarial network, which can be utilized to generate and optimize inorganic solid materials according to their properties. Reversible representation schemes for generative models are compared between molecular and crystalline structures, and challenges in regard to the latter are also discussed. Finally, we summarize the recent application of generative models in the exploration of chemical space with compositional and configurational degrees of freedom, and potential future directions are speculatively outlined.

## Keywords

Inverse design, inorganic solid materials, machine learning, generative model

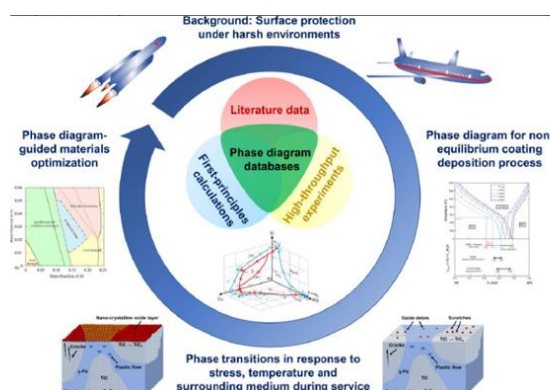
## Review

## 5. Development of robust surfaces for harsh service environments from the perspective of phase formation and transformation

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Lou, M.; Xu K.; Chen L.; Hong C.; Yuan Y.; Du Y.; Du Y.; Chang K. Development of robust surfaces for harsh service environments from the perspective of phase formation and transformation. *J. Mater. Inf.* 2021, 1, 5. <http://dx.doi.org/10.20517/jmi.2021.02>

### Abstract



The rise of the materials genome and materials informatics has enabled the accelerated development of robust surfaces for harsh service environments in the nuclear, aerospace and marine industries. Accurate information on the phase formation and transformation of materials (particularly coating materials) in synthesis and service processes is a prerequisite for the successful optimization of their properties. However, both these processes proceed under non-equilibrium conditions, making the traditional CALPHAD (CALculation of PHase Diagrams) approach incapable of describing the phase relation and stability. Hence, this study provides a brief review on the recent research advances pertaining to the phase formation during coating deposition, the phase transformation in service and the materials optimization targeted for demanding working conditions. We also summarize the challenges of expanding phase diagram databases with a wide adaptability to metastable phase formation and non-equilibrium phase transformation in multicomponent systems. Through the elaboration of each research case, this review provides new insights into the surface protection of materials serving in harsh environments.

### Keywords

Surface coating, harsh service environments, metastable phase formation, non-equilibrium phase transformation, phase diagram databases, CALPHAD, first-principles calculations, high-throughput experiments

### Review

## 6. Integrating computational materials science and materials informatics for the modeling of phase stability

[Full-Text](#) [PDF](#) [RIS](#) [Article Explanation Video](#)

**Cite this article:** Song, X.; Guo K.; Lu H.; Liu D.; Tang F. Integrating computational materials science and materials informatics for the modeling of phase stability. *J. Mater. Inf.* 2021, 1, 7. <http://dx.doi.org/10.20517/jmi.2021.06>

### Abstract

With rapid developments in big data and artificial intelligence technologies, materials informatics has become a new paradigm of materials science and engineering. In this review, the progress of modeling studies of phase stability in alloys is presented, with particular attention given to the development of the paradigm from traditional computational materials science (CMS) to materials informatics. The features of CMS models for phase stability studies are compared with those of data-driven approaches. The advantages of data-driven modeling in the framework of materials informatics are revealed. The approaches for developing interpretable machine learning, which has been mainly integrated with the developed CMS models and material science theories, are also discussed. Finally, the prospects for data-driven materials design based on the stability control of the dominant phases with regards to performance are proposed.

### Keywords

Phase stability, computational materials science, materials informatics, databases, machine learning, data-driven materials design

## Editorial

## 7. Informatics is fueling new materials discovery

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Zhang, T.Y.; Liu X.J. Informatics is fueling new materials discovery. *J. Mater. Inf.* 2021, 1, 6. <http://dx.doi.org/10.20517/jmi.2021.09>

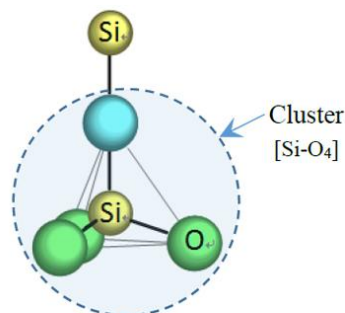
## Research Article

## 8. Composition genes in materials

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Zhang, S.; Wang Q.; Dong C. Composition genes in materials. *J. Mater. Inf.* 2021, 1, 8. <http://dx.doi.org/10.20517/jmi.2021.04>

### Abstract



Composition gene  $[\text{Si-O}_4]\text{Si}$  in  $\text{SiO}_2$

High-performance materials always possess specific chemical compositions. The present work points out that the composition genes, which are the basic structural units that serve as the composition carriers, are actually the molecule-like chemical units. Friedel oscillations, in combination with the cluster-plus-glue-atom model, are fully presented to show how to uncover the composition genes hidden in chemical short-range orders in any material. Examples are given in three categories of materials, i.e., metallic alloys including solid solutions and metallic glasses, inorganic compounds as well as relevant glasses, and polymers. Furthermore, materials can be classified into single-, dual-, and multi-gene types. The proposition of composition genes facilitates the understanding of prevailing materials and can be a useful tool to guide the exploration of new composition space.

### Keywords

Composition genes, chemical units, cluster-plus-glue-atom model, Friedel oscillations

## Review

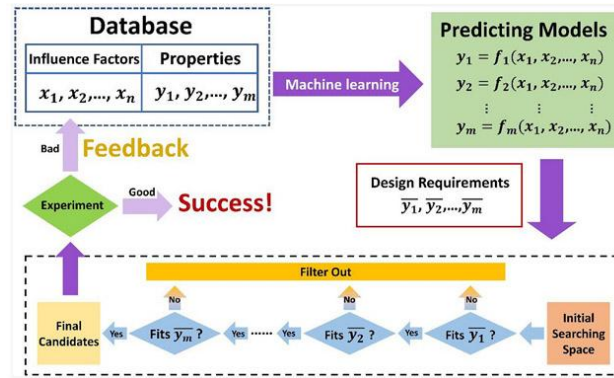
### 9. Machine learning-guided design and development of metallic structural materials

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Yu, J.; Xi S.; Pan S.; Wang Y.; Peng Q.; Shi R.; Wang C.; Liu X. Machine learning-guided design and development of metallic structural materials. *J. Mater. Inf.* 2021, 1, 9. <http://dx.doi.org/10.20517/jmi.2021.08>

### Abstract





In recent years, the advent of machine learning (ML) in materials science has provided a new tool for accelerating the design and discovery of new materials with a superior combination of mechanical properties for structural applications. In this review, we provide a brief overview of the current status of the ML-aided design and development of metallic alloys for structural applications, including high-performance copper alloys, nickel- and cobalt-based superalloys, titanium alloys for biomedical applications and high strength steel. We also present our perspectives regarding the further acceleration of data-driven discovery, development, design and deployment of metallic structural materials and the adoption of ML-based techniques in this endeavor.

### Keywords

Structural materials, metallic alloys, materials informatics, machine learning, composition-processing-microstructure-property relationships

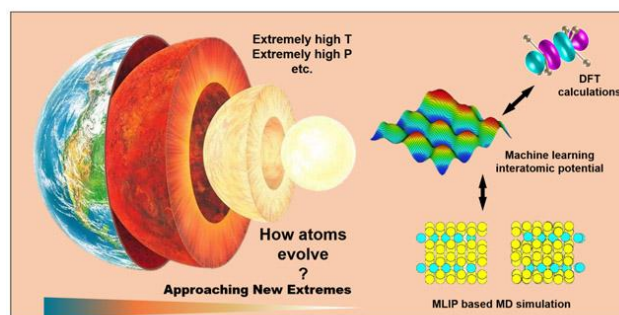
## Review

### 10. Taking materials dynamics to new extremes using machine learning interatomic potentials

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Yang, Y.; Zhao L.; Han C.X.; Ding X.D.; Lookman T.; Sun J.; Zong H.X. Taking materials dynamics to new extremes using machine learning interatomic potentials. *J. Mater. Inf.* 2021, 1, 10. <http://dx.doi.org/10.20517/jmi.2021.001>

### Abstract



Understanding materials dynamics under extreme conditions of pressure, temperature, and strain rate is a scientific quest that spans nearly a century. Atomic simulations have had a considerable impact on this endeavor because of their ability to uncover materials' microstructure evolution and properties at the scale of the relevant physical phenomena. However, this is still a challenge for most materials as it requires modeling large atomic systems (up to millions of particles) with improved accuracy. In many cases, the availability of sufficiently accurate but efficient interatomic potentials has become a serious bottleneck for performing these simulations as traditional potentials fail to represent the multitude of bonding. A new class of potentials has emerged recently, based on a different paradigm from the traditional approach. The new potentials are constructed by machine-learning with a high degree of fidelity from quantum-mechanical calculations. In this review, a brief introduction to the central ideas underlying machine learning interatomic potentials is given. In particular, the coupling of machine learning models with domain knowledge to improve accuracy, computational efficiency, and interpretability is highlighted. Subsequently, we demonstrate the effectiveness of the domain knowledge-based approach in certain select problems related to the kinetic response of warm dense materials. It is hoped that this review will inspire further advances in the understanding of matter under extreme conditions.

## Keywords

Machine learning, interatomic potential, extreme condition, domain knowledge, materials science

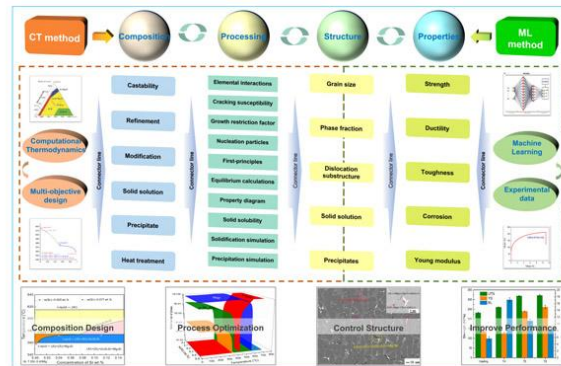
## Review

### 11. Boosting for concept design of casting aluminum alloys driven by combining computational thermodynamics and machine learning techniques

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Yi, W.; Liu G.; Gao J.; Zhang L. Boosting for concept design of casting aluminum alloys driven by combining computational thermodynamics and

## Abstract



Casting aluminum alloys are commonly used in industries due to their excellent comprehensive performance. Alloying/microalloying and post-solidification heat treatments are the most common measures to tune the microstructure for enhancing their mechanical properties. However, it is very challenging to achieve accurate and efficient development of novel casting aluminum alloys using the traditional trial-and-error method. With the rapid development of computer technology, the computational thermodynamics (CT) in the framework of the CALculation of PHase Diagram approach, the data-driven machine learning (ML) technique, and also their combinations have been proved to be effective approaches for the design of casting aluminum alloys. In this review, the state-of-the-art computational alloy design approaches driven by CT and ML techniques, as well as their combinations, were comprehensively summarized. The current status of the thermodynamic database for aluminum alloys, as the core for CT, was also briefly introduced. After that, a variety of successful case studies on the design of different casting aluminum alloys driven by CT, ML, and their combinations were demonstrated, including common applications, CT-driven design of Sc-additional Al-Si-Mg series casting alloys, and design of Srmodified A356 alloys driven by combining CT and ML. Finally, the conclusions of this review were drawn, and perspectives for boosting the computational design approach driven by combining CT and ML techniques were pointed out.

## Keywords

Casting aluminum alloy, alloy design, computational thermodynamics, machine learning

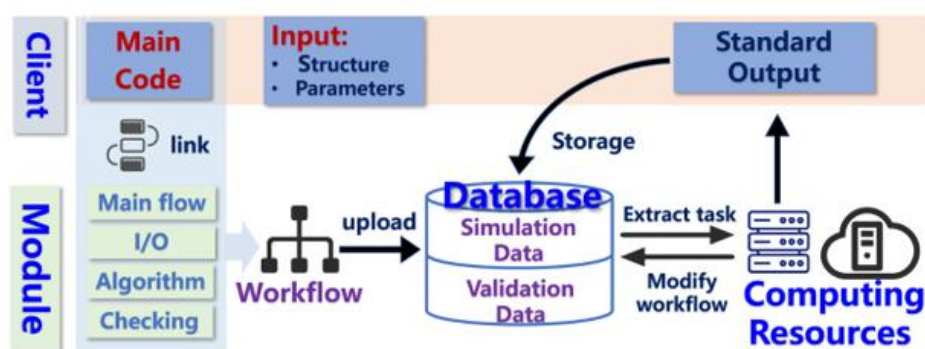
## Review

### 12. Big data-assisted digital twins for the smart design and manufacturing of advanced materials: from atoms to products

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Wang, W.Y.; Yin J.; Chai Z.; Chen X.; Zhao W.; Lu J.; Sun F.; Jia Q.; Gao X.; Tang B.; Hui X.; Song H.; Xue F.; Liu Z.K.; Li J. Big data-assisted digital twins for the smart design and manufacturing of advanced materials: from atoms to products. *J. Mater. Inf.* 2022, 2, 1. <http://dx.doi.org/10.20517/jmi.2021.11>

## Abstract



Motivated by the ever-increasing wealth of data boosted by national strategies in terms of data-driven Integrated Computational Materials Engineering (ICME), Materials Genome Engineering, Materials Genome Infrastructures, Industry 4.0, Materials 4.0 and so on, materials informatics represents a unique strategy in revealing the fundamental relationships in the development and manufacturing of advanced materials. Materials developments are becoming ever more integrated with robust data-driven and data-intensive technologies. In the present review, big data-assisted digital twins (DTs) for the smart design and manufacturing of advanced materials are presented from the perspective of the digital thread. In the introduction of the DT design paradigm in the ICME era, the simulation aspects of DT and the data and design infrastructures are discussed. Referring to the simulation and theoretical factors of DTs, high-throughput simulation and automation and artificial intelligence-assisted multiscale atomistic modeling are detailed through several cases studies. With respect to data and data mining technologies, entropy and its application for attribute selection in decision trees are discussed to emphasize knowledge-based modeling, simulation and data analysis in machine learning coherently. Guided by the perspectives and case studies of the digital thread, we present our recent work on the design, manufacturing and product service via big data-assisted DTs for smart design and manufacturing by integrating some of these advanced concepts and technologies. It is believed that big data-assisted DTs for smart design and manufacturing effectively support better products with the application of novel materials by reducing the time and cost of materials design and deployment.

## Keywords

ICME, Materials Genome Engineering, high-throughput, automation, workflow,

data mining, digital thread

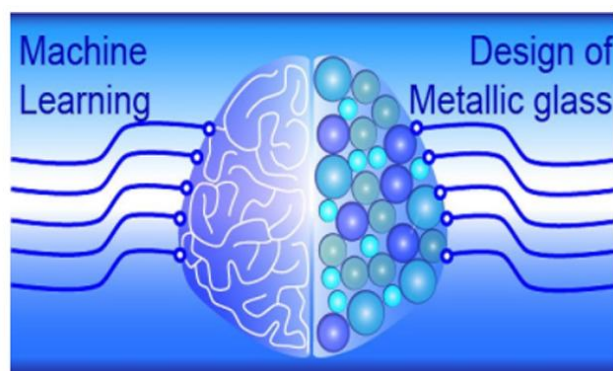
## Review

### 13. A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Zhou, Z.; Shang Y.; Yang Y. A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability. *J. Mater. Inf.* 2022, 2, 2. <http://dx.doi.org/10.20517/jmi.2021.12>

#### Abstract



The discovery of novel metallic glasses (MGs) with high glass-forming ability (GFA) has been an important area of active research for years in materials science and engineering. Unfortunately, the traditional approach based on trial-and-error methods is inefficient, time consuming and costly. Therefore, machine learning (ML) has recently drawn significant research interest as an alternative approach for the development of MGs. In this review, we discuss the current progress regarding the ML guided design of MGs from a variety of perspectives, including the GFA database, data representation, ML algorithms and numerical evaluation. Furthermore, we consider the challenges facing this field, including the scarcity and quality of GFA data, the development of physics informed data descriptors, the selection of appropriate algorithms and the necessity for experimental validation. We also briefly discuss possible solutions to tackle these challenges.

#### Keywords

Alloy design, metallic glasses, machine learning, glass-forming ability, data featurization

## Research Article

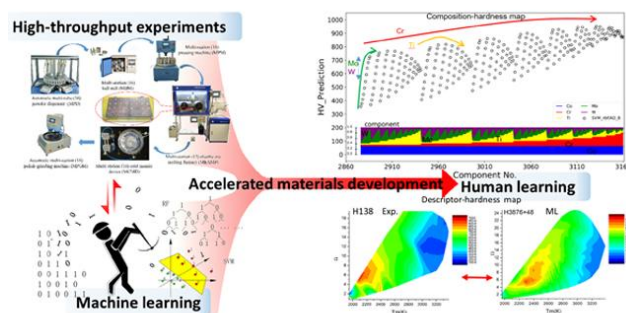
### 14. Accelerated development of hard high-entropy alloys with data-driven

## high-throughput experiments

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Liu, Y.; Wang J.; Xiao B.; Shu J. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments. *J. Mater. Inf.* 2022, 2, 3. <http://dx.doi.org/10.20517/jmi.2022.03>

## Abstract



The development of multicomponent alloys with target properties poses a significant challenge, owing to the enormous number of potential component combinations, high costs and the inefficiency of conventional empirical trial-and-error experimental approaches. To tackle this challenge, we develop a machine learning (ML)-guided high-throughput experimental (HTE) approach to accelerate the development of non-equimolar hard  $\text{Co}_x\text{Cr}_y\text{Ti}_z\text{Mo}_u\text{W}_v$  high-entropy alloys (HEAs). We first develop a set of all-process HTE facilities ranging from multi-tube ingredient assignment to multi-station electrical arc smelting and specimen preparation for bulk alloy samples with discrete compositions. Instead of random or combinatorial composition searching, HEAs with only  $\sim 1/28$  of all the potential compositions are synthesized in two stages guided by the ML prediction. The final ML models, trained using 138 experimental data, predict the alloy hardness with mean relative errors of 5.3%, 6.3% and 15.4% at high ( $\text{HV} > 800$ ), medium ( $\text{HV} = 600\text{--}800$ ) and low ( $\text{HV} < 600$ ) hardness ranges, respectively. In total, 14 superhard HEAs with  $\text{HV} > 900$  are discovered by our ML-guided HTE approach. Moreover, the multiple ML models predict the hardness of 3876 hypothetical alloys covering the whole composition range, thereby revealing the systematic component effects based on the complete composition-hardness and descriptor-hardness correlations. The hardening mechanisms are elaborated by analyzing the microstructures of  $\text{CoCrTiMoW}$ . Furthermore, physical insights can be gained by transitioning from “*machine learning*” to “*learning from machine*”. This work demonstrates that our ML-guided HTE approach provides an effective strategy for multicomponent alloy development with a potential hundred-fold overall increase in efficiency at a fraction of the cost compared to conventional methods.



High-throughput experiments, machine learning, multicomponent alloys, high-entropy alloys, hard alloys

## 15. Domain knowledge-guided interpretive machine learning: formula discovery for the oxidation behavior of ferritic-martensitic steels in supercritical water

**Cite this article:** Cao, B.; Yang S.; Sun A.; Dong Z.; Zhang T.Y. Domain knowledge-guided interpretive machine learning: formula discovery for the oxidation behavior of ferritic-martensitic steels in supercritical water. *J. Mater. Inf.* 2022, 2, 4. <http://dx.doi.org/10.20517/jmi.2022.04>

[illegible]

A general formula with high generalization and accurate prediction power is highly desirable for science, technology and engineering. In addition to human beings, artificial intelligence algorithms show great promise for the discovery of formulas. In this study, we propose a domain knowledge-guided interpretive machine learning strategy and demonstrate it by studying the oxidation behavior of ferritic-martensitic steels in supercritical water. The oxidation Cr equivalent is, for the first time, proposed in the present work to represent all contributions of alloying elements to oxidation, derived by our domain knowledge and interpretive machine learning algorithms. An open-source tree classifier for linear regression algorithm is also, for the first time, developed to materialize the formula with collected data. This algorithm effectively captures the linear correlation between compositions, testing environments and oxidation behaviors from the data. The sure independence screening and sparsifying operator algorithm finally assembles the information derived from the tree classifier for linear regression algorithm, resulting in a general formula. The general formula with the determined parameters has the power to predict, quantitatively and accurately, the oxidation behavior of ferritic-martensitic steels with multiple alloying elements exposed to

various supercritical water environments, thereby providing guidance for the design of anti-oxidation steels and hence promoting the development of power plants with improved safety. The present work demonstrates the power of domain knowledge-guided interpretive machine learning with respect to the data-driven discovery of physics-informed formulas and the acceleration of materials informatics development.

## Keywords

Domain knowledge, interpretive, oxidation Cr equivalent, tree classifier for linear regression (TCLR), general formula

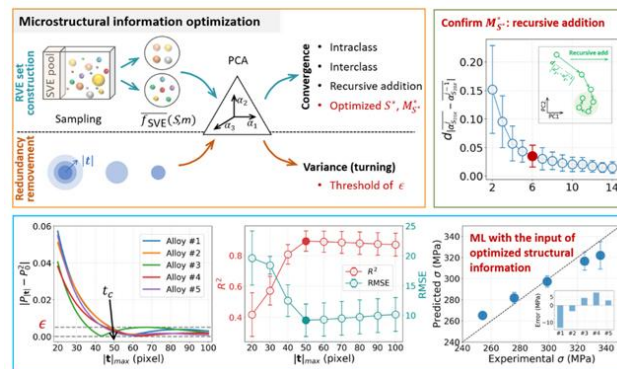
## Research Article

### 16. Structure-property modeling scheme based on optimized microstructural information by two-point statistics and principal component analysis

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Hu, X.; Zhao J.; Chen Y.; Wang Y.; Li J.; Wu Q.; Wang Z.; Wang J. Structure-property modeling scheme based on optimized microstructural information by two-point statistics and principal component analysis. *J. Mater. Inf.* 2022, 2, 5. <http://dx.doi.org/10.20517/jmi.2022.05>

## Abstract



Construction of the structure-property (SP) relationship is an important tenet during materials development. Optimizing microstructural information is a necessary and challenging task in understanding and improving this linkage. To solve the problem that the experimental microstructures with a small size usually fail to represent the entire sample structure, a data-driven scheme integrating two-point statistics, principal component analysis, and machine learning was developed to reasonably construct a representative volume element (RVE) set from the small microstructures and extract optimized structural information. Based on the elaborate quantitative metrics and method, this kind of RVE set was successfully constructed on an experimental microstructure dataset of ferrite



heat-resistant steels. Moreover, to remove redundant information included in two-point statistics, the critical threshold of the tolerance factor related to the coherence length in microstructures was determined to be 0.005. An accurate SP linkage was finally established (mean absolute error < 6.28 MPa for yield strength). This scheme was further validated on two other simulated and experimental datasets, which proved that it can offer scientific nature, reliability, and universality compared to traditional strategies. This scheme has a bright application prospect in microstructure classification, property prediction, and alloy design.

## Keywords

Microstructural information, structure-property linkage, two-point statistics, machine learning

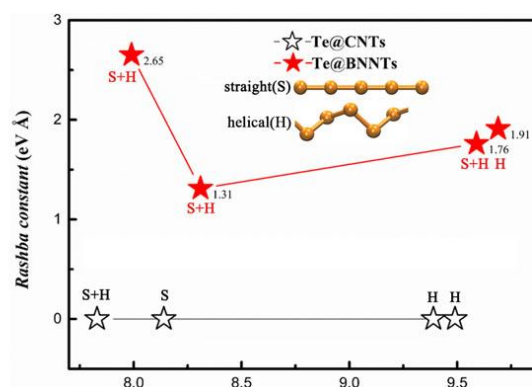
## Research Article

### 17. Giant rashba splitting of confined Te chains in nanotubes: the size-, chirality-, and type- effects of nanotubes

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Han, J.; Qi L.; Ma C.; Gao W. Giant rashba splitting of confined Te chains in nanotubes: the size-, chirality-, and type- effects of nanotubes. *J. Mater. Inf.* 2022, 2, 6. <http://dx.doi.org/10.20517/jmi.2022.08>

## Abstract



Understanding the coupling between one-dimensional (1D) materials and their protective materials is essential for developing nanodevices. Herein, we investigate the effect of the size, chirality, and type of nanotubes [such as carbon/boron nitride nanotubes (CNTs/BNNTs)] on the atomic and electronic structures of confined Te chains using density functional theory. We find that the optimal configurations of the Te chains confined in CNTs/BNNTs depend strongly on the size of the nanotubes but weakly on their chirality and type. Furthermore, the Te@BNNTs exhibit giant Rashba splitting with a Rashba constant of up to

2.65 eV Å, while the Te@CNTs show no splitting. This is mainly due to the large bandgap of the BNNTs, as well as the enhanced symmetry breaking of the Te chains when confined. Our findings provide a basis for the design of nano spin devices through protective materials.

## Keywords

Confined Te chains, Nanotubes, Size, Chirality, Rashba splitting

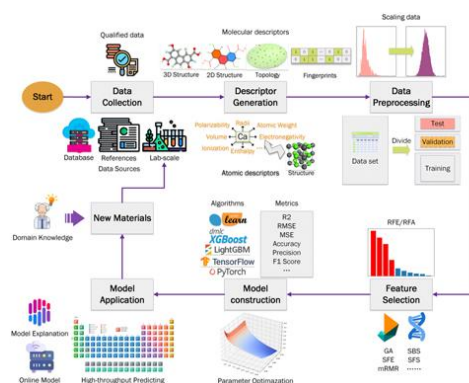
## Review

### 18. Recent progress in the data-driven discovery of novel photovoltaic materials

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Lu, T.; Li M.; Lu W.; Zhang T.Y. Recent progress in the data-driven discovery of novel photovoltaic materials. *J. Mater. Inf.* 2022, 2, 7. <http://dx.doi.org/10.20517/jmi.2022.07>

## Abstract



The discovery of new photovoltaic materials can facilitate technological progress in clean energy and hence benefit overall societal development. Machine learning (ML) and deep learning (DL) technologies integrated with domain knowledge are revolutionizing the traditional trial-and-error research paradigm that is associated with high costs, inefficiency, and significant human effort. This review provides an overview of the recent progress in the data-driven discovery of novel photovoltaic materials for perovskite, dye-sensitized and organic solar cells. The integral workflow of the ML/DL training progress is briefly introduced, covering data preparation, feature engineering, model building and their applications. The cutting-edge challenges and issues in the ML/DL workflow are summarized specifically for photovoltaic materials. Real examples are emphasized to illustrate how to utilize ML/DL techniques in the discovery of novel photovoltaic materials. The prospects and future directions of the data-driven discovery of novel photovoltaic materials are also provided.

## Keywords

Machine learning, materials design, deep learning, photovoltaic materials, data-driven, perovskite solar cells, organic solar cells, dye-sensitized solar cells

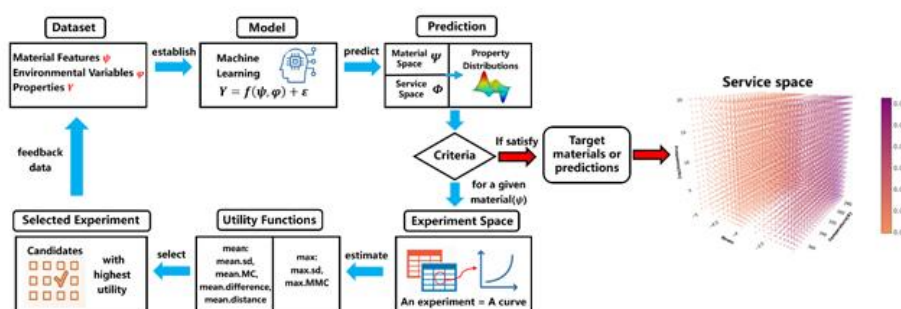
## Research Article

### 19. Estimating the performance of a material in its service space via Bayesian active learning: a case study of the damping capacity of Mg alloys

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Shi, B.; Zhou Y.; Fang D.; Tian Y.; Ding X.; Sun J.; Lookman T.; Xue D. Estimating the performance of a material in its service space via Bayesian active learning: a case study of the damping capacity of Mg alloys. *J. Mater. Inf.* 2022, 2, 8. <http://dx.doi.org/10.20517/jmi.2022.06>

#### Abstract



In addition to being determined by its chemical composition and processing conditions, the performance of a material is also affected by the variables of its service space, including temperature, pressure, and frequency. A rapid means to estimate the performance of a material in its service space is urgently required to accelerate the screening of materials with targeted performance. In the present study, a materials informatics approach is proposed to rapidly predict performance within a service space based on existing data. We utilize an active learning loop, which employs an ensemble machine learning method to predict the performance, followed by a Bayesian experimental design to minimize the number of experiments for refinement and validation. This approach is demonstrated by predicting the damping properties of a ZE62 magnesium alloy in a service space defined by frequency, strain amplitude, and temperature based on the available data for other magnesium alloys. Several utility functions that recommend a particular experiment to refine the estimates of the service space are used and compared. In particular, the standard deviation is found to reduce the prediction error most efficiently. After augmenting the database with nine new experimental measurements, the uncertainties associated with the predicted damping capacities are largely reduced. Our method allows us to forecast the properties in the service

space of a given material by rapid refinement of the predictions via experiment measurements.

### Keywords

Ensemble learning, active learning, Mg alloys, damping, service space, Bayesian optimization

## Research Article

### 20. High-cycle fatigue S-N curve prediction of steels based on a transfer learning-guided convolutional neural network

[Full-Text](#) [PDF](#) [RIS](#) [Article Explanation Video](#)

**Cite this article:** Wei, X.; Wang C.; Jia Z.; Xu W. High-cycle fatigue S-N curve prediction of steels based on a transfer learning-guided convolutional neural network. *J. Mater. Inf.* 2022, 2, 9. <http://dx.doi.org/10.20517/jmi.2022.12>

### Abstract

The evaluation and prediction of fatigue properties are crucial for metallic materials. Although the determination of S-N curves represents the most important methods for evaluating such properties, its fatigue testing is costly and time-consuming. Furthermore, fatigue testing involves different test conditions, thereby complicating the evaluation of the fatigue properties. This study develops a transfer convolutional neural network (TR-CNN) framework, in which the prediction of the reversed torsion S-N curves of steels is transferred from rotating bending S-N curves. In the TR-CNN framework, the source CNN models for rotating-bending curve prediction are first trained based on the composition and process conditions. Subsequently, based on the source models, the reversed torsion S-N curves are estimated by training the TR-CNN models based on only a small dataset. After proving the rationality of the framework, its universality with respect to different amounts of data is further investigated. The reversed torsion curves under small-sample conditions (22 samples) are predicted accurately by the TR-CNN. Additionally, the TR-CNN models remain accurate under varying amounts of data (22-112 samples), showing excellent generality for different amounts of fatigue data. The predictive capability of the TR-CNN models is improved by introducing tensile properties into the source models. The proposed TR-CNN framework can significantly reduce the cost of evaluating fatigue properties, and the prediction of S-N curves can be optimized by combining the transfer framework and low-cost properties related to fatigue.

### Keywords

High-cycle fatigue, S-N curves, CNN, transfer learning

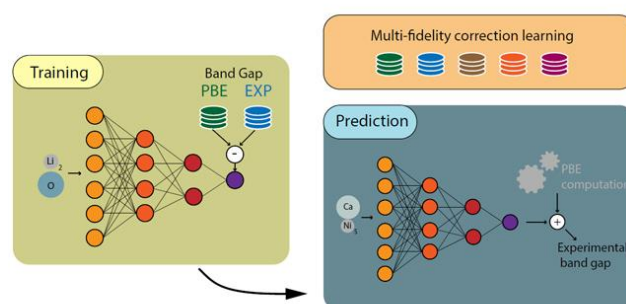
## Research Article

### 21. Accurate experimental band gap predictions with multifidelity correction learning

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** De Breuck, P.P.; Heymans G.; Rignanese G.M. Accurate experimental band gap predictions with multifidelity correction learning. *J. Mater. Inf.* 2022, 2, 10. <http://dx.doi.org/10.20517/jmi.2022.13>

#### Abstract



To improve the precision of machine-learning predictions, we investigate various techniques that combine multiple quality sources for the same property. In particular, focusing on the electronic band gap, we aim at having the lowest error by taking advantage of all available experimental measurements and density-functional theory calculations. We show that learning about the difference between high- and low-quality values, considered a correction, significantly improves the results compared to learning on the sole high-quality experimental data. As a preliminary step, we also introduce an extension of the MODNet model, which consists of using a genetic algorithm for hyperparameter optimization. Thanks to this, MODNet is shown to achieve excellent performance on the Matbench test suite.

#### Keywords

Machine learning, electronic band gap, multi-fidelity, transfer-learning, materials properties

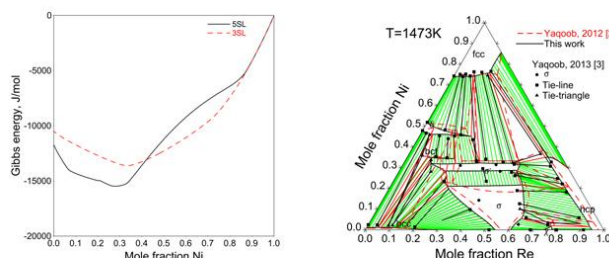
## Research Article

### 22. A thermodynamic database of the Ni-Mo-Re system

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Lei, X.H.; Liu W.; Luo F.; Lu X.G. A thermodynamic database of the Ni-Mo-Re system. *J. Mater. Inf.* 2022, 2, 11.

## Abstract



Thermodynamic databases are essential prerequisites for developing advanced materials, such as Ni-based superalloys. The present work collects a large amount of experimental and first-principles calculation data concerning the thermodynamics and phase diagrams of the Ni-Mo-Re system, based on which the thermodynamic properties of the ternary and its binary sub-systems Ni-Mo and Mo-Re are assessed by means of the CALculation of PHase Diagrams (CALPHAD) approach. The thermodynamic database containing all model parameters is established and most experimental data are reproduced satisfactorily. The present work demonstrates the use of the CALPHAD method as a practical appliance in the toolbox of materials informatics to analyze and discriminate various types of data by thermodynamic modeling and then produce more useful data in wider ranges of compositions and temperatures by computational predictions.

## Keywords

Ni-Mo-Re, first-principles calculations, thermodynamic modeling, CALPHAD

## Viewpoints

### 23. Additive manufacturing as a tool for high-throughput experimentation

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Xiong, W. Additive manufacturing as a tool for high-throughput experimentation. *J. Mater. Inf.* 2022, 2, 12. <http://dx.doi.org/10.20517/jmi.2022.19>

## Abstract

Additive manufacturing (AM) is a disruptive technology with a unique capability in fabricating parts with complex geometry and fixing broken supply chains. However, many AM techniques are complicated with their processing features due to complex heating and cooling cycles with the melting of feedstock materials.

Therefore, it is quite challenging to directly apply the materials design and processing optimization method used for conventional manufacturing to AM techniques. In this viewpoint paper, we discuss some of the ongoing efforts of high-throughput (HT) experimentation, which can be used for materials development and processing design. Particularly, we focus on the beam- and powder-based AM techniques since these methods have demonstrated success in HT experimentation. In addition, we propose new opportunities to apply AM techniques as the materials informatic tools contributing to materials genome.

### Keywords

Additive manufacturing, integrated computational materials engineering, materials genome, materials informatics, machine learning

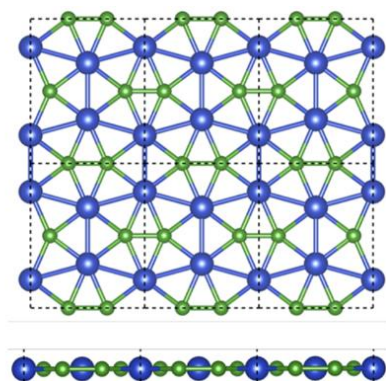
## Research Article

### 24. CuB monolayer: a novel 2D anti-van't Hoff/Le Bel nanostructure with planar hyper-coordinate boron/copper and superconductivity

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Tu, K.; Gu J.; Lu L.; Yuan S.; Zhou L.; Chen Z. CuB monolayer: a novel 2D anti-van 't Hoff/Le Bel nanostructure with planar hyper-coordinate boron/copper and superconductivity. *J. Mater. Inf.* 2022, 2, 13. <http://dx.doi.org/10.20517/jmi.2022.10>

### Abstract



To achieve specific applications, it is always desirable to design new materials with peculiar topological properties. Herein, based on a  $D_{2h}$   $B_2Cu_6H_6$  molecule with the unique chemical bonding of planar pentacoordinate boron (ppB) as a building block, we constructed an infinite CuB monolayer by linking  $B_2Cu_6$  subunits in an orthorhombic lattice. The planarity of the CuB sheet is attributed to the multicenter bonds and electron donation-back donation, as revealed by chemical bonding analysis. As a global minimum confirmed by the particle swarm



optimization method, the CuB monolayer is expected to be highly stable, as indicated by its rather high cohesive energy, absence of soft phonon modes, and good resistance to high temperature, and thus is highly feasible for experimental realization. Remarkably, this CuB monolayer is metallic and predicted to be superconducting with an estimated critical temperature ( $T_c$ ) of 4.6 K, and the critical temperature could be further enhanced by tensile strains (to 21 K at atmospheric pressure).

## Keywords

Planar pentacoordinate boron, planar heptacoordinate copper, copper boride monolayer, density functional calculations, superconducting

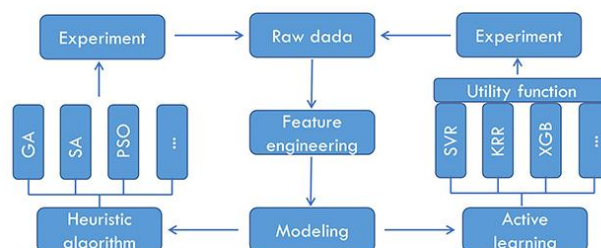
## Review

### 25. A mini review of machine learning in inorganic phosphors

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Jiang, L.; Jiang X.; Lv G.; Su Y. A mini review of machine learning in inorganic phosphors. *J. Mater. Inf.* 2022, 2, 14. <http://dx.doi.org/10.20517/jmi.2022.21>

## Abstract



Machine learning has promoted the rapid development of materials science. In this review, we provide an overview of recent advances in machine learning for inorganic phosphors. We take two aspects of material properties prediction and optimization based on iterative experiments as entry points to outline the applications of machine learning for inorganic phosphors in terms of Debye temperature prediction and luminescence intensity and thermal stability optimization. By analyzing the machine learning methods and their application objectives, current problems are summarized and suggestions for subsequent development are proposed.

## Keywords



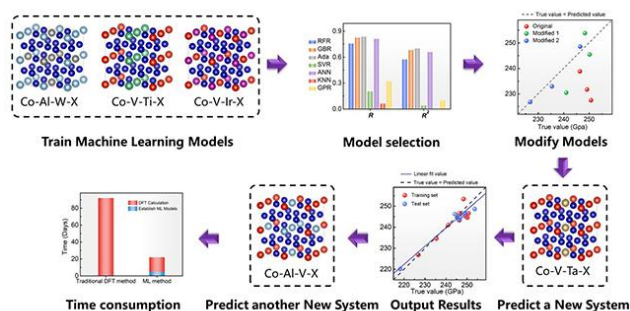
## Research Article

26. Machine learning-accelerated first-principles predictions of the stability and mechanical properties of L<sub>12</sub>-strengthened cobalt-based superalloys

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Xi, S.; Yu J.; Bao L.; Chen L.; Li Z.; Shi R.; Wang C.; Liu X. Machine learning-accelerated first-principles predictions of the stability and mechanical properties of L<sub>12</sub>-strengthened cobalt-based superalloys. *J. Mater. Inf.* 2022, 2, 15. <http://dx.doi.org/10.20517/jmi.2022.22>

## Abstract



As promising next-generation candidates for applications in aero-engines, L<sub>12</sub>-strengthened cobalt (Co)-based superalloys have attracted extensive attention. However, the L<sub>12</sub> strengthening phase in first-generation Co-Al-W-based superalloys is metastable, and both its solvus temperature and mechanical properties still need improvement. Therefore, it is necessary to discover new L<sub>12</sub>-strengthened Co-based superalloy systems with a stable L<sub>12</sub> phase by exploring the effect of alloying elements on their stability. Traditional first-principles calculations are capable of providing the crystal structure and mechanical properties of the L<sub>12</sub> phase doped by transition metals but suffer from low efficiency and relatively high computational costs. The present study combines machine learning (ML) with first-principles calculations to accelerate crystal structure and mechanical property predictions, with the latter providing both the training and validation datasets. Three ML models are established and trained to predict the occupancy of alloying elements in the supercell and the stability and mechanical properties of the L<sub>12</sub> phase. The ML predictions are evaluated using first-principles calculations and the accompanying data are used to further refine the ML models. Our ML-accelerated first-principles calculation approach offers more efficient predictions of the crystal structure and mechanical properties for Co-V-Ta- and Co-Al-V-based systems than the traditional

counterpart. This approach is applicable to expediting crystal structure and mechanical property calculations and thus the design and discovery of other advanced materials beyond Co-based superalloys.

## Keywords

Co-based superalloys, first-principles calculations, site occupancy, phase stability, mechanical properties, machine learning

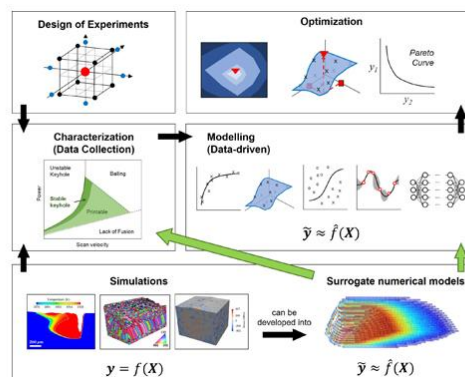
## Review

### 27. Process parameter optimization of metal additive manufacturing: a review and outlook

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Chia, HY.; Wu J.; Wang X.; Yan W. Process parameter optimization of metal additive manufacturing: a review and outlook. *J. Mater. Inf.* 2022, 2, 16. <http://dx.doi.org/10.20517/jmi.2022.18>

## Abstract



The selection of appropriate process parameters is crucial in metal additive manufacturing (AM) as it directly influences the defect formation and microstructure of the printed part. Over the past decade, research efforts have been devoted to identifying "optimal" processing regimes for different materials to achieve defect-free manufacturing, which mostly involve costly trial-and-error experiments and computationally expensive mechanistic simulations. Hence, it is apropos to critically review the methods used to achieve the optimal process parameters in AM. This work seeks to provide a structured analysis of current methodologies and discuss systematic approaches toward general optimization work in AM and the process parameter optimization of new AM alloys. A brief review of process-induced defects due to process parameter selection is given and the current methods for identifying "optimal processing windows" are summarized. Research works are analyzed under a standard optimization framework, including the design of experiments and characterization, modelling

and optimization algorithms. The research gaps that preclude multi-objective optimization in AM are identified and future directions toward optimization work in AM are discussed. With growing capabilities in AM, we should reconsider the definition of the "optimal processing region".

## Keywords

Additive manufacturing, process parameters, optimization, modeling, design of experiments

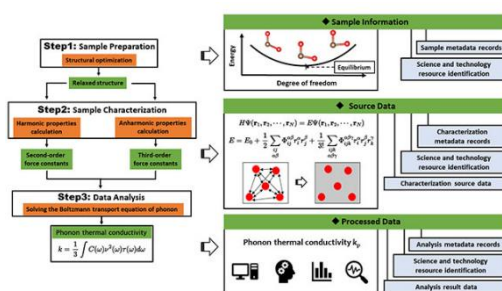
## Research Article

### 28. A metadata schema for lattice thermal conductivity from first-principles calculations

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Rao, Y.; Lu Y.; Zhang L.; Ju S.; Yu N.; Zhang A.; Chen L.; Wang H. A metadata schema for lattice thermal conductivity from first-principles calculations. *J. Mater. Inf.* 2022, 2, 17. <http://dx.doi.org/10.20517/jmi.2022.20>

## Abstract



Materials genome engineering databases represent fundamental infrastructures for data-driven materials design, in which the data resources should satisfy the FAIR (Findable, Accessible, Interoperable and Reusable) principles. However, a variety of challenges, such as data standardization, veracity and longevity, still impede the progress of data-driven materials science, including both high-throughput experiments and simulations. In this work, we propose a metadata schema for lattice thermal conductivity from first-principles calculations. The calculation workflow for lattice thermal conductivity includes structural optimization and the calculation of interatomic force constants and lattice thermal conductivity. The data generated during the calculation process corresponds to the virtual sample information, virtual source data and processed data, respectively, as specified in the *General rule for materials genome engineering data* of the Chinese Society for Testing and Materials. Following this general rule, the metadata structure and schema for each action are systematically defined and all metadata elements can be collected completely. Although this metadata schema is specific to lattice

thermal conductivity calculations, it provides general rules and insights for other computational materials data in materials genome engineering.

## Keywords

Materials genome engineering, metadata schema, first-principles calculations, lattice thermal conductivity

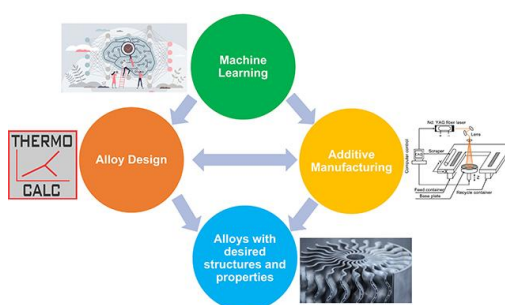
## Review

### 29. New trends in additive manufacturing of high-entropy alloys and alloy design by machine learning: from single-phase to multiphase systems

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Zhou, Y.; Zhang Z.; Wang D.; Xiao W.; Ju J.; Liu S.; Xiao B.; Yan M.; Yang T. New trends in additive manufacturing of high-entropy alloys and alloy design by machine learning: from single-phase to multiphase systems. *J. Mater. Inf.* 2022, 2, 18. <http://dx.doi.org/10.20517/jmi.2022.27>

## Abstract



Alloys with excellent properties are always in significant demand for meeting the severe conditions of industrial applications. However, the design strategies of traditional alloys based on a single principal element have reached their limits in terms of property optimization. The concept of high-entropy alloys (HEAs) provides a new design strategy based on multicomponent elements, which may overcome the bottleneck problems that exist in traditional alloys. To further maximize the capability of HEAs, a novel additive manufacturing (AM) technique has been utilized to produce HEA components with the desired structures and properties. This review considers a new trend in the AM of HEAs, i.e., from the AM of single-phase HEAs to multiphase HEAs. Although most as-printed single-phase HEAs show superior tensile properties to as-cast ones, their strength is still not satisfactory, especially at elevated temperatures. Thus, multiphase HEAs are developed by introducing hard second phases, such as  $L_{12}$ , BCC, carbides, oxides, nitrides, and so on. These phases can be introduced to the matrix using in situ alloying during AM or the subsequent heat treatment. Dislocation strengthening is considered as the main reason for improving the tensile properties

of as-printed single-phase HEAs. In contrast, multiple strengthening and toughening mechanisms occur in as-printed multiphase HEAs, which can synergistically enhance their mechanical properties. Furthermore, machine learning provides an effective method to design new alloys with the desired properties and predict the optimal AM parameters for the designed alloys without tedious experiments. The synergistic combination of machine learning and AM will significantly speed up scientific advances and promote industrial applications.

## Keywords

High-entropy alloys, additive manufacturing, precipitation hardening, strengthening mechanism, machine learning

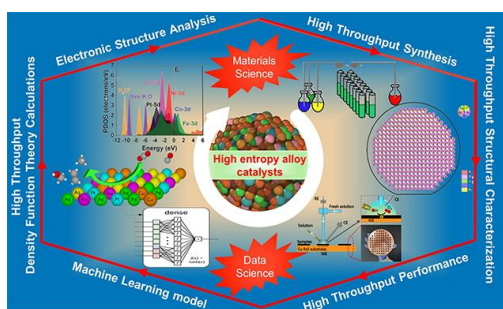
## Review

### 30. High-entropy alloy catalysts: high-throughput and machine learning-driven design

[Full-Text](#) [PDF](#) [RIS](#)

**Cite this article:** Chen, L.; Chen Z.; Yao X.; Su B.; Chen W.; Pang X.; Kim K.S.; Singh CV.; Zou Y. High-entropy alloy catalysts: high-throughput and machine learning-driven design. *J. Mater. Inf.* 2022, 2, 19. <http://dx.doi.org/10.20517/jmi.2022.23>

## Abstract



High-entropy alloy (HEA) catalysts have recently attracted worldwide research interest due to their promising catalytic performance. Most current studies focus on designing HEA catalysts through trial-and-error methods. This produces scattered data and is not conducive to obtaining a fundamental understanding of the structure-property-performance relationships for HEA catalysts, thereby hindering their rational design. High-throughput (HT) techniques and machine learning (ML) methods show significant potential in generating, processing and analyzing databases with a vast amount of data, providing a new strategy for the further development of HEA catalysts. In this review, we summarize the recent literature on HT techniques for HEA synthesis, characterization and performance testing. We also review the ML models that are used to process and analyze

existing databases to accelerate the discovery of HEA catalysts. Finally, the potential challenges and perspectives of HT techniques and ML models are presented to accelerate the discovery of new HEA catalysts and promote their development.

### **Keywords**

High-entropy alloys, catalysts, high-throughput, machine learning, structure-activity relationship