Science
AAAS

Supplementary Materials for

Machine learning–enabled high-entropy alloy discovery

Ziyuan Rao *et al*.

Corresponding authors: Ye Wei, y.wei@mpie.de; Dierk Raabe, d.raabe@mpie.de

Science **378**, 78 (2022)
DOI: 10.1126/science.abo4940

The PDF file includes:

Materials and Methods

HEA Generative Alloy Design (HEA-GAD). As shown in fig. S6 the continuous latent composition representation is learned by WAE in an unsupervised manner. WAE is one variant of the VAE family (*21*), which is defined as a constrained autoencoder architecture designed for unsupervised learning. The basis for WAE training is the optimization of an objective function (loss function) which takes into account a reconstruction loss and a regularization loss term. The loss function can be mathematically expressed as:

$$L = L_{rec} + L_{reg}$$

In the current scheme, L_{rec} is defined as binary cross-entropy of input-output pair and L_{reg} is defined as maximum mean discrepancy (MMD) (*48*) between the latent space distribution and a Gaussian prior. MMD is a kernel-based statistical measure used to determine whether given two

distributions are the same. Two distributions are similar if their moments are similar. The kernel transforms the latent space and prior distribution such that all moments are computed. In such a way, one can compute the difference between the moments, which gives a measure of similarity between latent space and the prior. Such a difference is defined as MMD. Mathematically, it can be expressed as:

$$MMD(P, Q) = \| E_{X\sim P}[\varphi(X)] - E_{Y\sim Q}[\varphi(Y)] \|_H$$

where P and Q are two distributions over a set S. The MMD is defined by a *feature map* φ: $S \rightarrow H$, H is defined as the reproducing kernel Hilbert space. Such loss function can effectively regularize the encoder variance as described in (*49*). We used an inverse multiquadric kernel for measuring the MMD, which is a common choice for such a task (*50*). Sigma is 8.0 and MMD_lambda is set to be 7.5×10^{-5}. The encoder block of WAE consists of 3 neural layers, and the number of nodes is 80, 64, and 48, respectively. Each layer is followed by layer normalization. The activation function is Leaky REctified Linear Unit (RELU) with a slope of -0.02. The decoder block is the same architecture as the encoder but arranged in reverse order. The model architecture is kept relatively simple to avoid overfitting, the main focus of hyperparameters tuning includes the number of epochs, batch size, and learning rate. The optimal setting is 400, 6, and 2.5×10^{-4}, respectively. Such settings were found by 10,000 random searches followed by 10 rounds of Bayesian Optimization (BO).

We used a GMM $Q_\xi(z)$ to estimate the density in the latent z space (i.e., the marginal posterior $q_\varphi(z)$). GMM is a density estimation model that fits the data points using a mixture of a finite number of Gaussian distributions with unknown parameters. The number of Gaussian distributions is usually determined via the empirical elbow method. The elbow method is

a heuristic used in determining the potential optimal number of clusters in a dataset. In our case, the average negative log-likelihood was plotted as a function of the number of clusters and we selected the 'elbow' of the curve as the number of Gaussian distributions. The elbow plot is shown in Supplementary fig. S7. In addition, an independent property classifier $q_\xi(c_{low-TEC}|z)$ was trained to fit the encoded compositions with labels (z_i, $c_{low-TEC}$), where compositions with thermal expansion coefficient (TEC) below 6.0×10^{-6} K^{-1} were labeled as low-TEC. The binary classifier is an MLP with one hidden layer and was trained with 5-fold cross-validation. The resulting accuracy is 90% and the training history is provided in Supplementary fig. S8.

In HEA-GAD, the conditioned generation of new synthetic compositions with low-TEC, i.e., $\boldsymbol{z} \sim p(\boldsymbol{z}|\boldsymbol{c}_{low-TEC})$, was realized using Metropolis-Hastings MCMC sampling in the latent $\boldsymbol{z}$ space.

$$p(\boldsymbol{z}|\boldsymbol{c}_{low-TEC}) \approx \tilde{p}_\xi(\boldsymbol{z}|\boldsymbol{c}_{low-TEC})$$

$$= \frac{Q_\xi(\boldsymbol{z})\, q_\xi(\boldsymbol{c}_{ow-TEC}|\boldsymbol{z})}{q_\xi(\boldsymbol{c}_{low-TEC})}$$

where $\tilde{p}_\xi(\boldsymbol{z}|\boldsymbol{c}_{low-TEC})$ is the approximation of $p(\boldsymbol{z}|\boldsymbol{c}_{low-TEC})$ using the GMM $Q_\xi(\boldsymbol{z})$ and the binary property classifier $q_\xi(\boldsymbol{c}_{low-TEC}|\boldsymbol{z})$, and $q_\xi(\boldsymbol{c}_{low-TEC})$ is the normalizing constant. The sampling of the MCMC algorithm begins with an initial composition in the latent space $z^{i=0} \in \boldsymbol{z}$. At each time index *i*, a new composition candidate $z' \in z$ will be drawn from the proposal distribution $q(z'|z^{i-1})$, where z^{i-1} represents the composition at time index *i*-1. Then, the proposed candidate z' will be either accepted $z^i = z'$ or rejected $z^i = z^{i-1}$ according to the acceptance probability $A(z^{i-1}, z')$. Here, we use a symmetric Gaussian distribution as the proposal distribution $q(z'|z)$. As a result, the acceptance probability is calculated as follow:

$$A(z, z') = \min\left\{1, \frac{\tilde{p}_\xi(z'|\boldsymbol{c}_{low-TEC})\, q(z|z')}{\tilde{p}_\xi(z|\boldsymbol{c}_{low-TEC})\, q(z'|z)}\right\}$$

$$= \min\left\{1, \frac{Q_\xi(z')\, q_\xi(\boldsymbol{c}_{ow-TEC}|z')}{Q_\xi(z)\, q_\xi(\boldsymbol{c}_{ow-TEC}|z)}\right\}$$

Two-stage ensemble regression model. The data preprocessing procedures are as follows:

- Atomic properties were normalized.
- Stratified 5-fold cross-validation was introduced to assure that data were sampled based on their distribution.
- 85:15 training testing split.

The mean absolute percentile error is selected as the loss function, it can be expressed as:

$$\text{MAPE Loss} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{A_i - P_i}{A_i}\right|$$

where A_i is the actual value, P_i is the prediction and n is the number of data. Since it assigns more weights to the data with lower TEC values. It suits better to the Invar prediction problem than the Mean Squared Error (MSE), as we are looking for a lower TEC value.

The first and second stages of TERM share the same regression model architecture but differ in their input features. The first stage's input is defined as compositions with the corresponding atomic features (see composition+atomic features in table S3). The model architecture consists of two basic blocks: the NN block and GBDT. NN block consists of 50 different NN models. Such 50 models were selected through 10, 000 iterations of random search and 10 iterations of Bayesian search. The hyperparameters to be determined are the number of layers, number of nodes, *etc*. Moreover, the bias induced by the training testing dataset split affects

the model accuracy significantly, due to the small size of the current dataset. Therefore the 50 models were retrained with 10 different random seeds to minimize this bias. GBDT block was constructed in a similar way using random search, Bayesian search and different random seeds. Therefore, it contains 50 different models as well. To accelerate the speed and avoid repeated predictions, we avoid calculating the physical properties of all the compositions at one time. Instead, we choose 10-30 compositions with affinity propagation (shown in Supplementary figs. S9-10) for the DFT and thermodynamic calculations.

The second stage of TERM aims at predicting the TEC of the top 3 compositions from the first stage. Firstly, the Density-Functional Theory (DFT) and thermodynamic calculations of the whole dataset were prepared. Secondly, the ensemble model architecture was trained in the same way as the first stage with this enhanced dataset. Finally, the top 3 compositions were chosen based on the ranking score. The BO in the current work uses Gaussian Process Regression (GPR) as the surrogate model. The rank of the hyperparameters by BO is illustrated in tables S5-8.

Exploration and exploitation strategy. Uncertainty is another helpful indicator in the search for Invar compositions. Uncertainty measures the "distance" of an unexplored composition to the known dataset. The larger the "distance", the higher the uncertainty which will be attributed to this composition. The exploration and exploitation strategy is a central concept in active learning (*51, 52*). It is a trade-off scheme aimed at gradual improvement towards a certain goal in the face of an unknown environment. In such a scheme, actions are guided by an objective function. In the high-entropy Invar scenario, the goal is to find the compositions that have the lowest possible TEC, and mathematically it can be written as an optimization problem:

$$\text{find input } \boldsymbol{x} \in \boldsymbol{H}$$

to the function f: $\boldsymbol{H}\rightarrow\boldsymbol{R}$

such that $\boldsymbol{x} = \arg\min_{\boldsymbol{x}\in\boldsymbol{H}} \{f(\boldsymbol{x})\}$

H is the compositional space, *f* is the function that maps the composition and its related properties to TEC and *R* is the space spanned all possible TEC values. The choice of *x* is guided by the mean and uncertainty of the TEC values. Exploration prefers the composition with higher uncertainty, while exploitation favors the composition with lower predicted TEC. However, such a baseline scheme faces two challenges in our case: firstly, even if the TERM training and testing results show good training and testing accuracy, the experimental TEC can still deviate significantly from the TERM prediction (even outside the uncertainty range, see Table 1). Secondly, while the uncertainty does reflect their 'distance' to the known data, we often observe that the experimental values often deviate from the predictions. Such deviation is due to the highly nonlinear nature of the composition-property relation and the small dataset (see Supplementary fig. S2). Hence, the value-based objective function cannot reliably guide the exploitation-exploration scheme. To overcome such issues, we proposed a rank-based objective function as a guide to our experiment. The rank of data is a very effective and widely used transformation of data (*32, 33*). By introducing the ranking as a metric instead of an actual prediction, the impact of outliers is mitigated and the inhomogeneity or skewness in the predicted TEC distribution eliminated, because all ranks are equally far apart from each other. More importantly, the rank-based objective function ensures that the candidate selection is less affected by model inaccuracy, and it provides a systematic way to combine model prediction and uncertainty by assigning different weights on the rank (instead of value) of TEC prediction and uncertainty:

Rank score = $\alpha \times$ rank_of_prediction + $\beta \times$ rank_of_uncertainty, with $\alpha + \beta = 1$

The rank-based objective function ensures that the candidate selection is less affected by model inaccuracy. Overall, our TERM exploration-exploitation scheme assigns the same weights to prediction and uncertainty (α = 0.9 and β = 0.1) with a combination of exploration and exploitation. The candidate selection is based on their descending order of rank score (see tables S9-14).

DFT calculations. The DFT calculations for face-centered cubic (FCC) Fe-Co-Ni-Cr-V-Cu alloys were performed with the exact muffin-tin orbital (EMTO) method (*53*) with the generalized gradient approximation for the exchange-correlation functional (*54*). We treated chemical disorder with the coherent potential approximation (*55*) and modeled the paramagnetic state with the disordered local moment configuration (*56*). The adopted total number of k-points in the irreducible Brillouin zone was 6281. A list of computed alloys compositions can be found in table S15. For each concentration, we computed 6 energy-volume points spanning an interval of ± 3% from the equilibrium volume. For all the DFT calculations, we have checked and assured that there is no abrupt change of local magnetic moment on each element. More specifically, while searching for the equilibrium volume, we avoided covering the lower volume range which may contain the low-spin to high-spin state transition especially for Fe. For HEAs containing Fe at the FCC phase, the high-spin state of Fe which gives lower energy than its low-spin state (*40*) was assured. We fitted the energy-volume curves with the Birch-Murnaghan equation of state (*57*) and derived the equilibrium energy E_0 and the equilibrium volume V_0 for the ferromagnetic and paramagnetic states, from which we estimated the magnetostriction ω_s. To obtain a smooth mapping between composition and property, we interpolated the concentration dependence of ω_s with the k-nearest neighbors (kNN) algorithm (*58*). The results are shown in table S15.

The finite-temperature lattice parameters of A2 and A3 alloys were obtained by fitting the equation of states (EOS) with the Helmholtz free energy described by a sum of the vibrational term, which is based on a quasi-harmonic Debye-Grüneisen model (*59*), and the contribution from magnetic entropy derived from the mean-field expression $S_{mag}/k_B = 2\eta \sum_{i=\mathrm{Fe,\ Ni,\ Co,\ Cr}} x_i \ln(1+\mu_i)$ (*60*). Here, x_i is the atomic concentration of different components and μ_i is the corresponding local magnetic moment. η denotes the pseudo-alloy concentration ($0 \le \eta \le 0.50$). In specific, to simulate the reduced-magnetization state of the A2 and A3 alloys at a finite temperature, the A2 and A3 alloys are represented by $(\mathrm{Fe}^{\uparrow}_{1-\eta}\mathrm{Fe}^{\downarrow}_{\eta})_{50.1}(\mathrm{Ni}^{\uparrow}_{1-\eta}\mathrm{Ni}^{\downarrow}_{\eta})_{16.7}(\mathrm{Co}^{\uparrow}_{1-\eta}\mathrm{Co}^{\downarrow}_{\eta})_{26.1}(\mathrm{Cr}^{\downarrow}_{1-\eta}\mathrm{Cr}^{\uparrow}_{\eta})_{7.1}$ and $(\mathrm{Fe}^{\uparrow}_{1-\eta}\mathrm{Fe}^{\downarrow}_{\eta})_{42.7}(\mathrm{Ni}^{\uparrow}_{1-\eta}\mathrm{Ni}^{\downarrow}_{\eta})_{9.1}(\mathrm{Co}^{\uparrow}_{1-\eta}\mathrm{Co}^{\downarrow}_{\eta})_{39.5}(\mathrm{Cr}^{\downarrow}_{1-\eta}\mathrm{Cr}^{\uparrow}_{\eta})_{8.7}$, respectively. The η value was acquired according to the magnetization versus temperature curve calculated using the Monte Carlo (MC) method implementing the UppASD code *(57)*. The required magnetic exchange coupling parameters by MC simulations were calculated by employing the generalized perturbation method (GPM) *(58)*. implementing the local density approximation (LDA) exchange-correlation functional.

Thermodynamic calculations. The CALPHAD calculations were performed with ThermoCalc 2019b (*61*) using the High Entropy Alloys v.2.1.1 database on the same compositional grid as the DFT calculations. For each composition, we computed the equilibrium phase fractions and compositions of the BCC and FCC phases. ω_s and T_{C} were then calculated using the kNN mapping at the predicted equilibrium compositions for fcc and bcc.

Experimental sample preparation. The designed alloys were prepared in a mini-arc furnace (hearing power 5.3 kW, 50/60 Hz, argon atmosphere of 800 bar), and solidified into cylindrical samples with 4 mm diameter by suction casting. The chemical compositions of the alloys were

measured by inductively coupled plasma (ICP) mass spectrometry (table S16). Then we used the same heat treatment shown in the papers (*34,35*) where we got the training data: the alloys were annealed at 1000 ºC for 1 hour and slowly cooled with the furnace.

Microstructure characterization and thermal expansion curves measurement. Before the microstructural analysis, fine polishing of the samples' surfaces was performed using an oxide suspension (OPS) with silica particle sizes around 50 nm for more than 30 min. Electron backscatter diffraction (EBSD) measurements were measured on a Zeiss-Crossbeam XB 1540 FIB scanning electron microscope (SEM) with a Hikari camera and the TSL OIM data collection software. Back-scattered electron imaging (BSEI) was performed on a Zeiss-Merlin SEM. The elemental distributions were analyzed using energy-dispersive X-ray spectroscopy (EDS) at the microscale. EBSD map, BSE image and corresponding EDS maps of the $Fe_{48.8}Ni_{17.8}Co_{22.2}Cr_{6.2}Cu_5$ (wt.%) HEA (B2 in Table 1) and the reference alloy $Fe_{63.5}Ni_{36.5}$ (wt.%) HEA are shown in figs. S11-12.

The thermal expansion curves were measured in a PPMS equipped with a dilatometer option from 5 K to 400 K with a heating speed of 1 K/min. The cubic samples with dimensions of 2×2×3 mm^3 were first ground using silicon carbide paper from 600 to 4000 granulation. Experimental results of the thermal expansion behavior of the alloys from 1st to 6th rounds are shown in fig. S13.

The samples listed in table S3 are prepared and measured with the same methods we mentioned above.

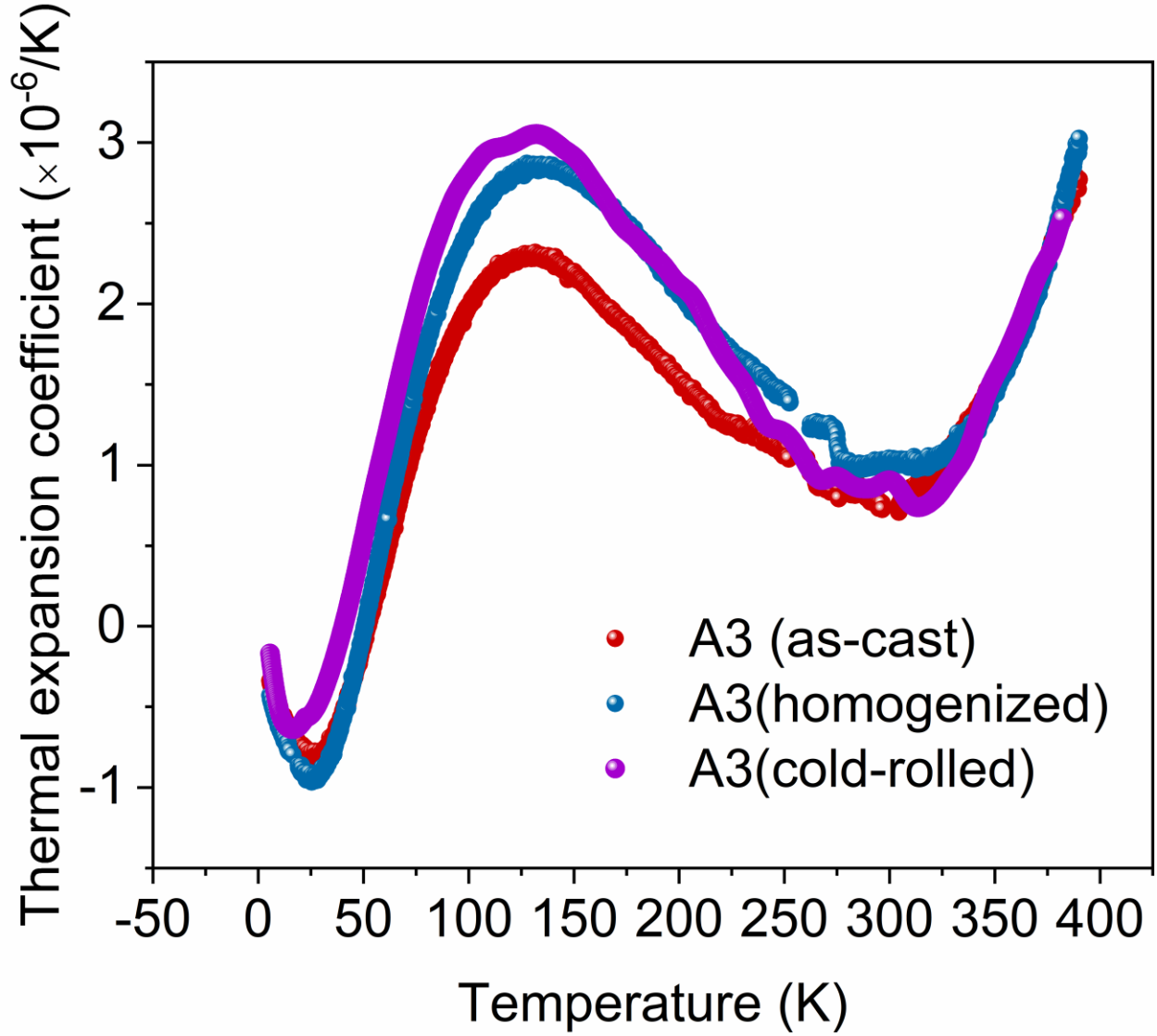

Figure S1. The microstructures and TECs of $Fe_{41.84}Ni_{9.37}Co_{40.85}Cr_{7.94}$ (wt.%) (A3 in Table 1) with different treatments. The as-cast, homogenized and 50% cold-rolled alloys share similar thermal expansion behavior. The TEC difference between them is less than 1×10^{-6}/ºC around room temperature.

Initial dataset: ~700 (696 from literature + 3 from previous work)

Experiment speed: 2 weeks 3 new experiments performed

Training

Alloys	Numbers	Number of possible compositions (step size 0.1%)	Percentage of known compositions in all possible compositions
Fe-Ni	19	1001	≈1.90%
Fe-Co	17	1001	≈1.69%
Ni-Co	13	1001	≈1.30%
Fe-Ni-Co	69	501502	≈0.014%
Fe-Co-Cr	93	501502	≈0.019%
Fe-Co-Cr-Cu	32	167668502	$\approx 1.9\times10^{-7}$
Fe-Co-Cr-Ni	291	167668502	$\approx 1.7\times10^{-6}$
Fe-Co-Ni-V	162	167668502	$\approx 9.7\times10^{-7}$

Prediction

Alloys	Numbers	Number of possible compositions (step size 0.1%)	Percentage of known compositions in all possible compositions
Fe-Co-Cr-Ni	291	167668502	$\approx 1.7\times10^{-6}$
Fe-Co-Cr-Ni-Cu	3	42084793752	$\approx 7.1\times10^{-11}$

Figure S2. The initial dataset and experimental speed. 696 data points are taken from the publications (*34-39*). 3 data points are taken from our previous work shown in table S3. For training, the initial dataset only occupies a very small percentage of all possible compositions. For prediction, we have 291 quaternary and 3 quinary HEAs which make the prediction even more difficult. The step size of the composition 0.1% is considered as the resolution of the arc-melting furnace casting.

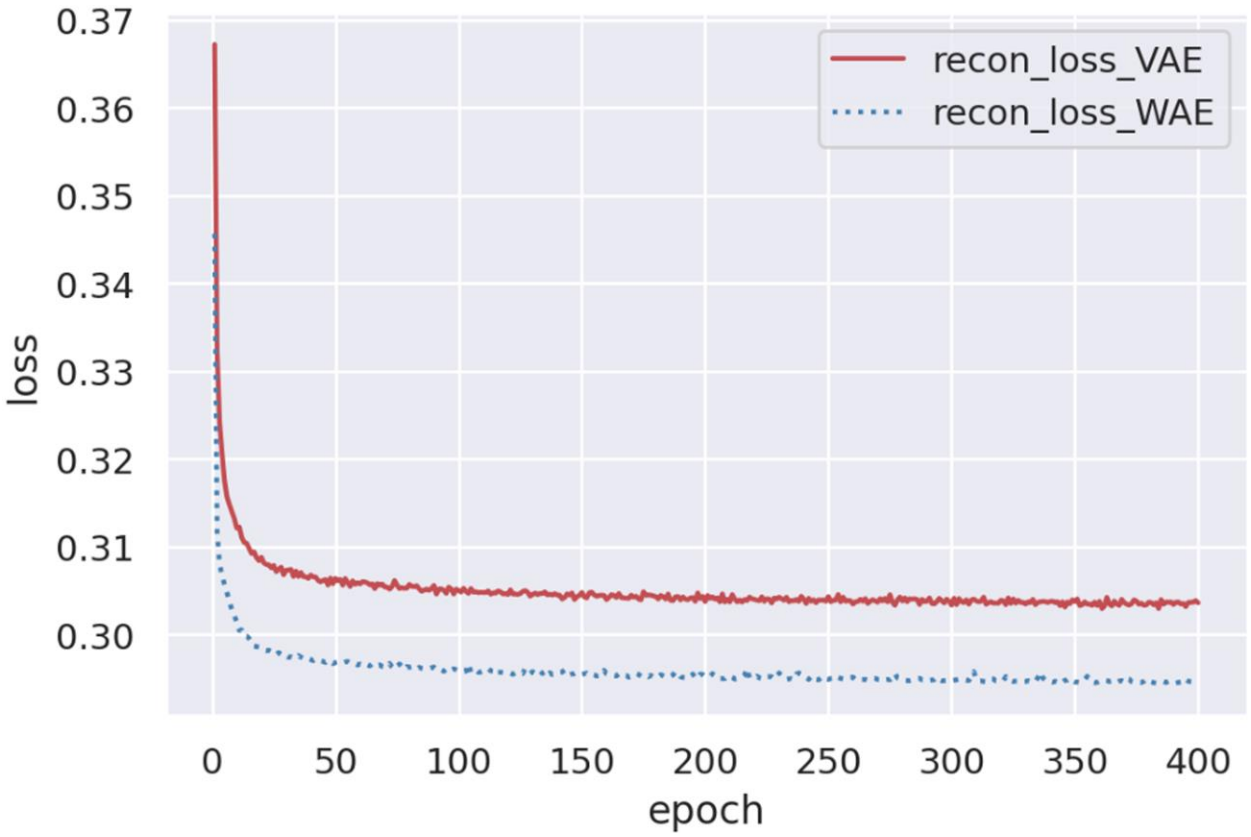

Figure S3. The comparison of wasserstein autoencoder (WAE) and variational autoencoder (VAE) as a generative model. WAE has a lower reconstruction loss of the whole data compared with VAE. Both algorithms use a batch size of 6, a leaning rate of $2.5e^{-4}$, and an epoch number of 400. The hyperparameters are optimized with Bayesian optimization.

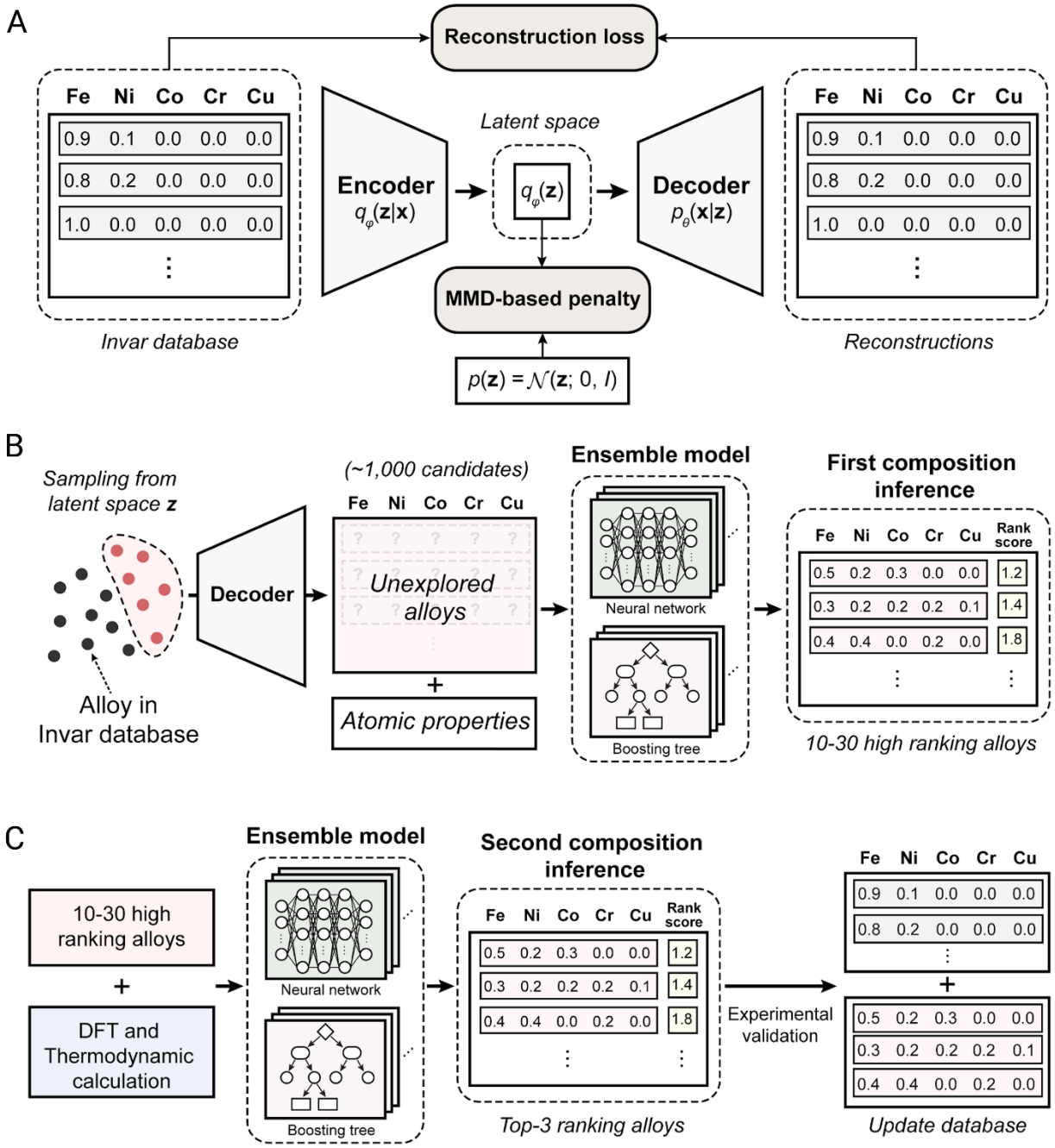

Figure S4. The schematic framework of the three modeling steps. (a) A HEA composition generating scheme (HEA-GAD) based on a deep generative model. The encoder $q_\varphi(z|x)$ with parameters φ takes compositions of alloys as input and the decoder $p_\theta(x|z)$ with parameters θ can act as a generator for suggesting new alloy compositions based on the learned latent z

representation. (b) A large-scale screening with multilayer neural network model and gradient boosting decision tree. The new synthetic compositions with potentially low-TEC from $p(\boldsymbol{z}|\boldsymbol{c}_{low-TEC})$ are sampled through the Markov Chain Monte Carlo method based on the latent z space shown in (a). The screening utilizes the mean TEC prediction value from 100 models (including 50 multilayer perceptrons and 50 gradient boosting decision trees) as the final prediction, while the corresponding variance as the uncertainty. A rank-based objective function works as a guide to the final choice. (c) Physically informed screening model with the integration of DFT and thermodynamic calculation. The TEC of the top 3 candidates will be experimentally measured by PPMS and the results will be fed back to the training database for the next active learning iteration.

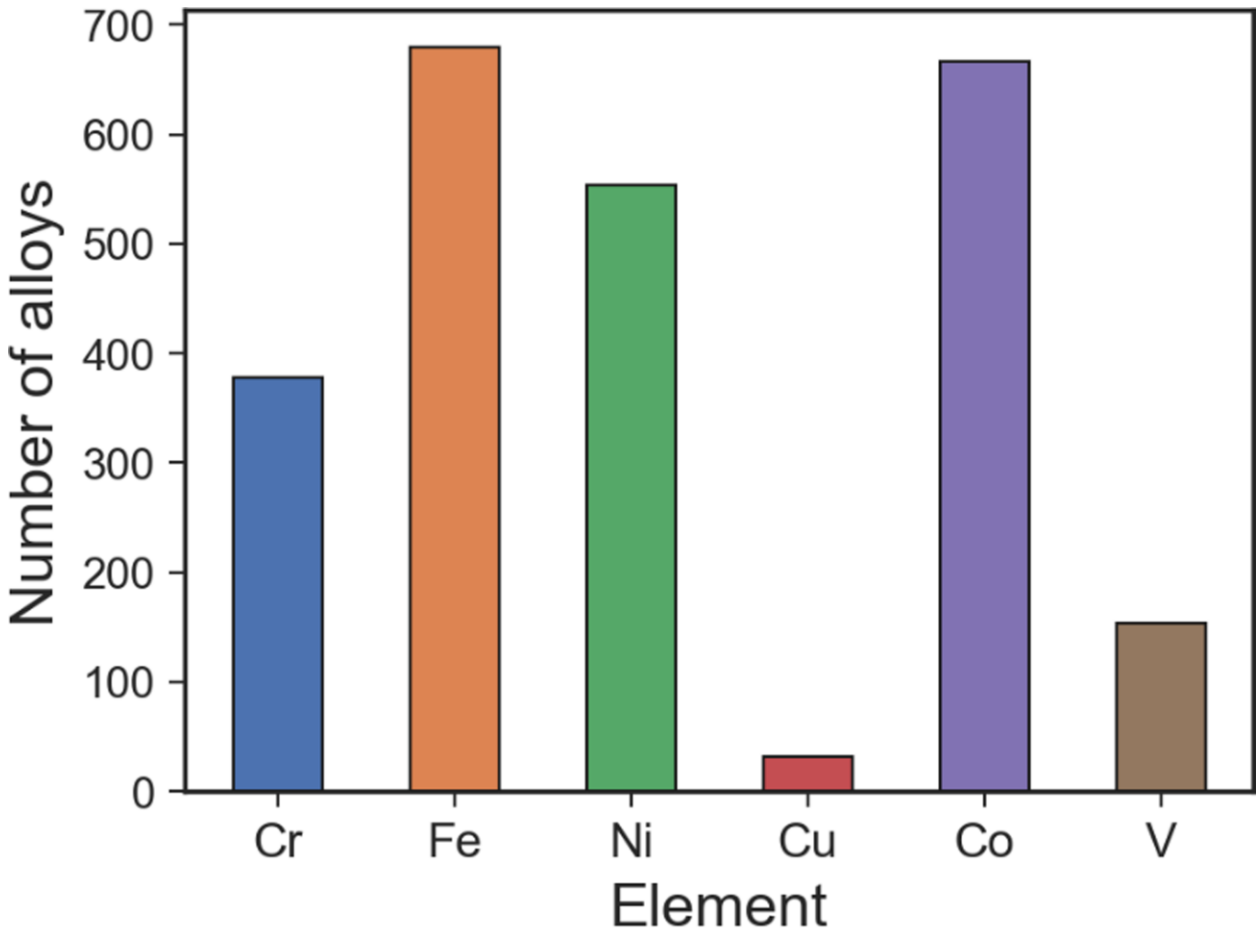

Figure S5. Numbers of alloys with different elements in the dataset. The vast majority of alloy in the dataset doesn't contain Cu at all and only a few alloys have 5% Cu.

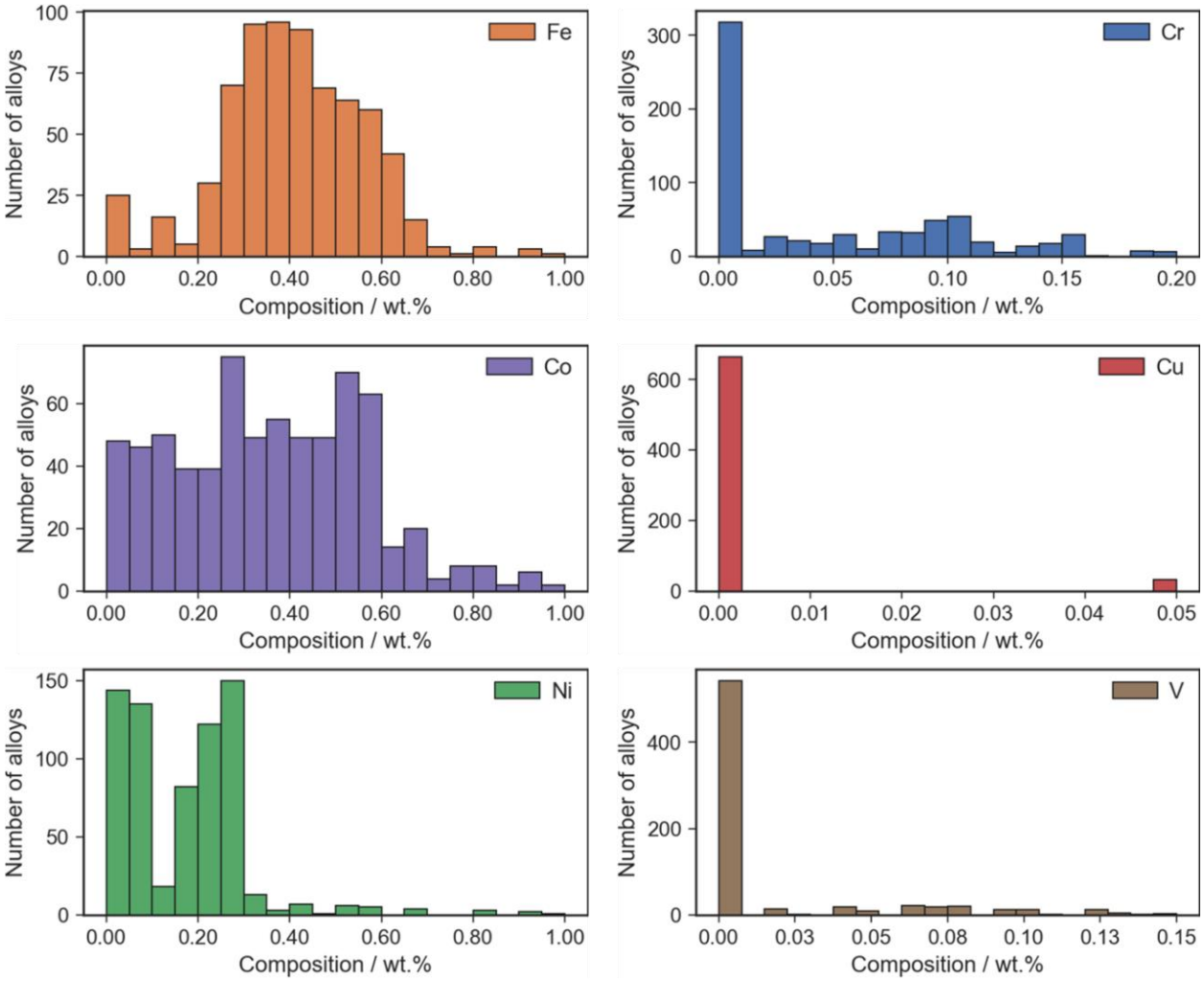

Figure S6. Distribution of different elements in the dataset. Cu has an extremely imbalanced distribution and other elements have various distributions.

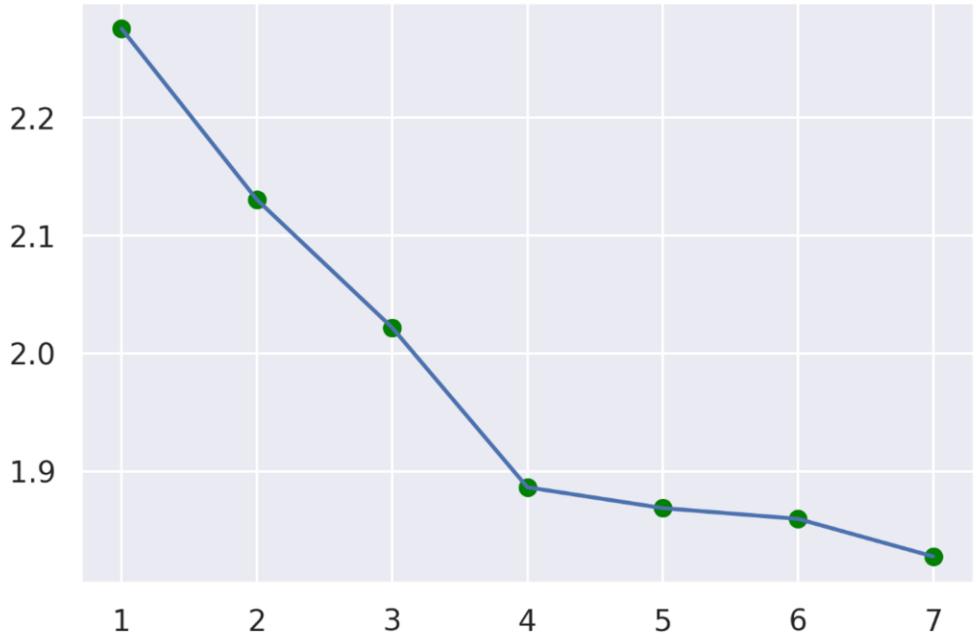

Figure S7. The elbow plot of the GMM training determining the potential optimal number of clusters.

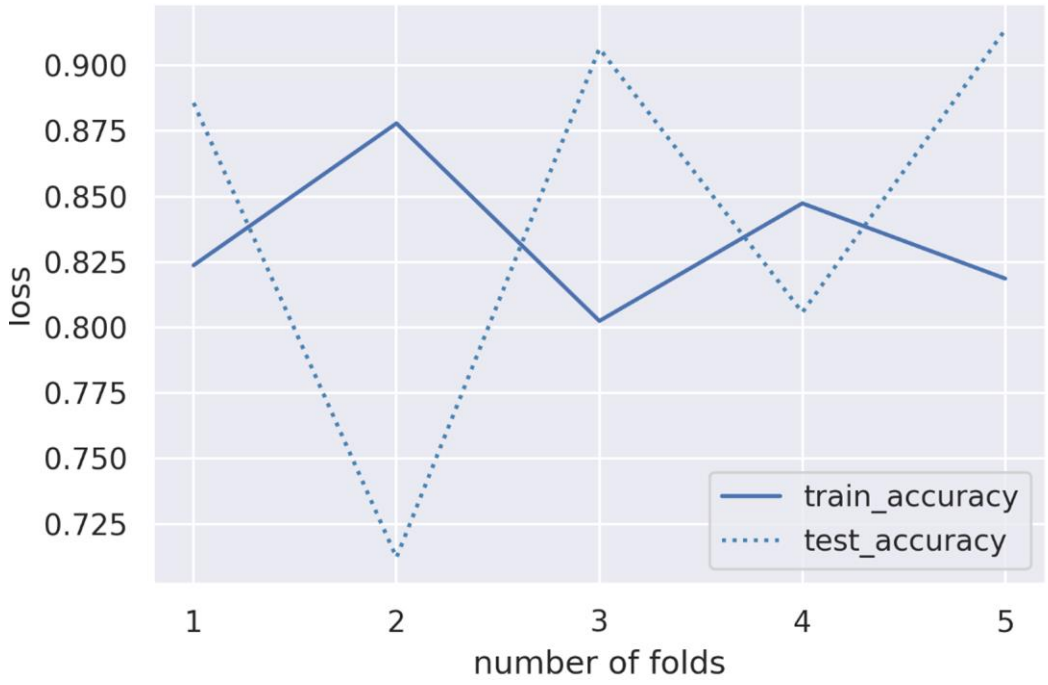

Figure S8. The training results of the neural network binary classifier. We use 5 fold cross-validation (CV) and 100 epochs for each CV. The batch size is 16 and the learning rate is 1×10^{-4}.

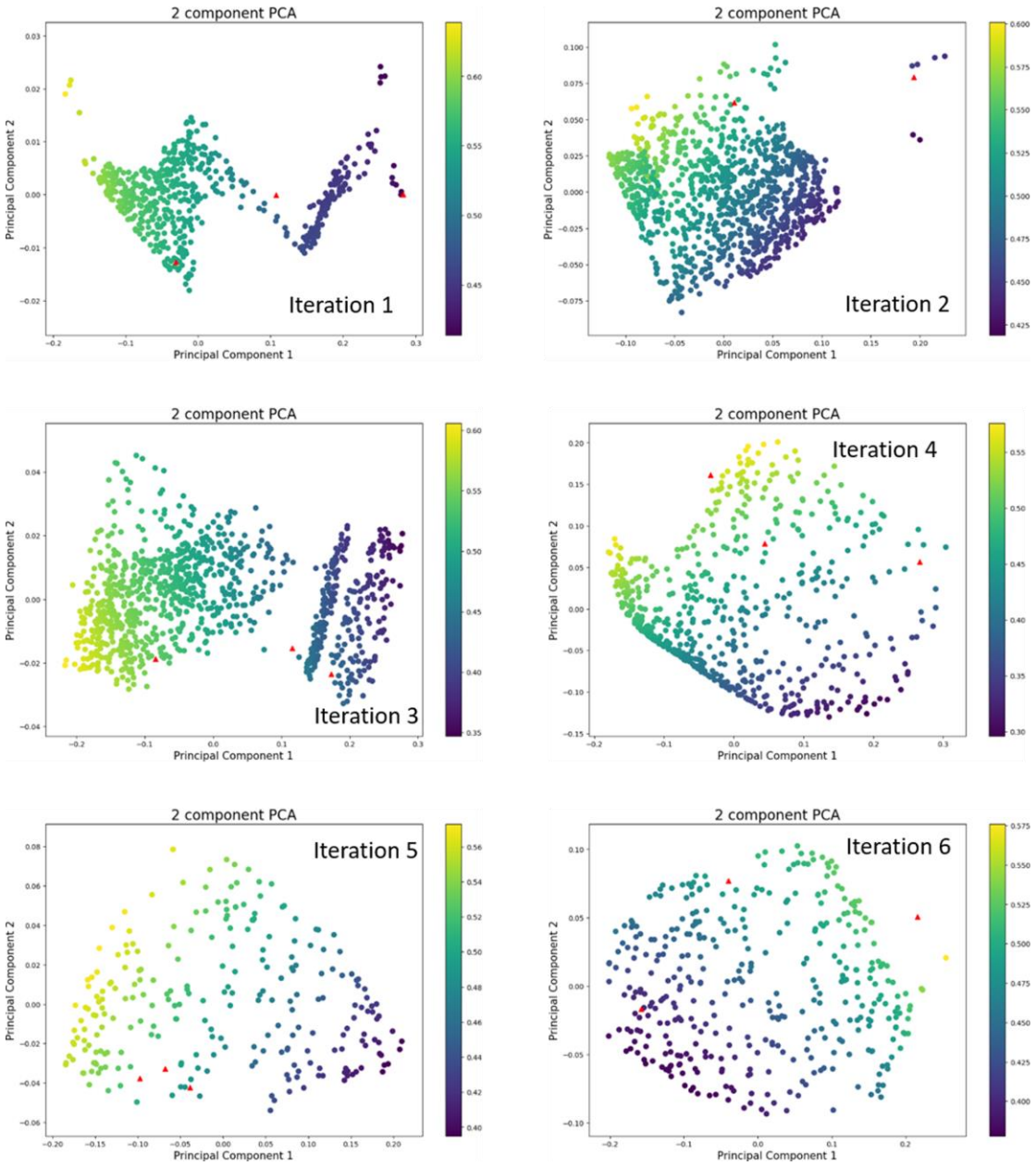

Figure S9. Sampled compositions from the first stage of the Two-stage Ensemble Regression Model (TERM). Latent space of the samples for iterations 1 to 6 generated by principal component analysis (PCA). The color changes with the change of Fe composition. The red triangles are the alloys prepared for experimental measurements.

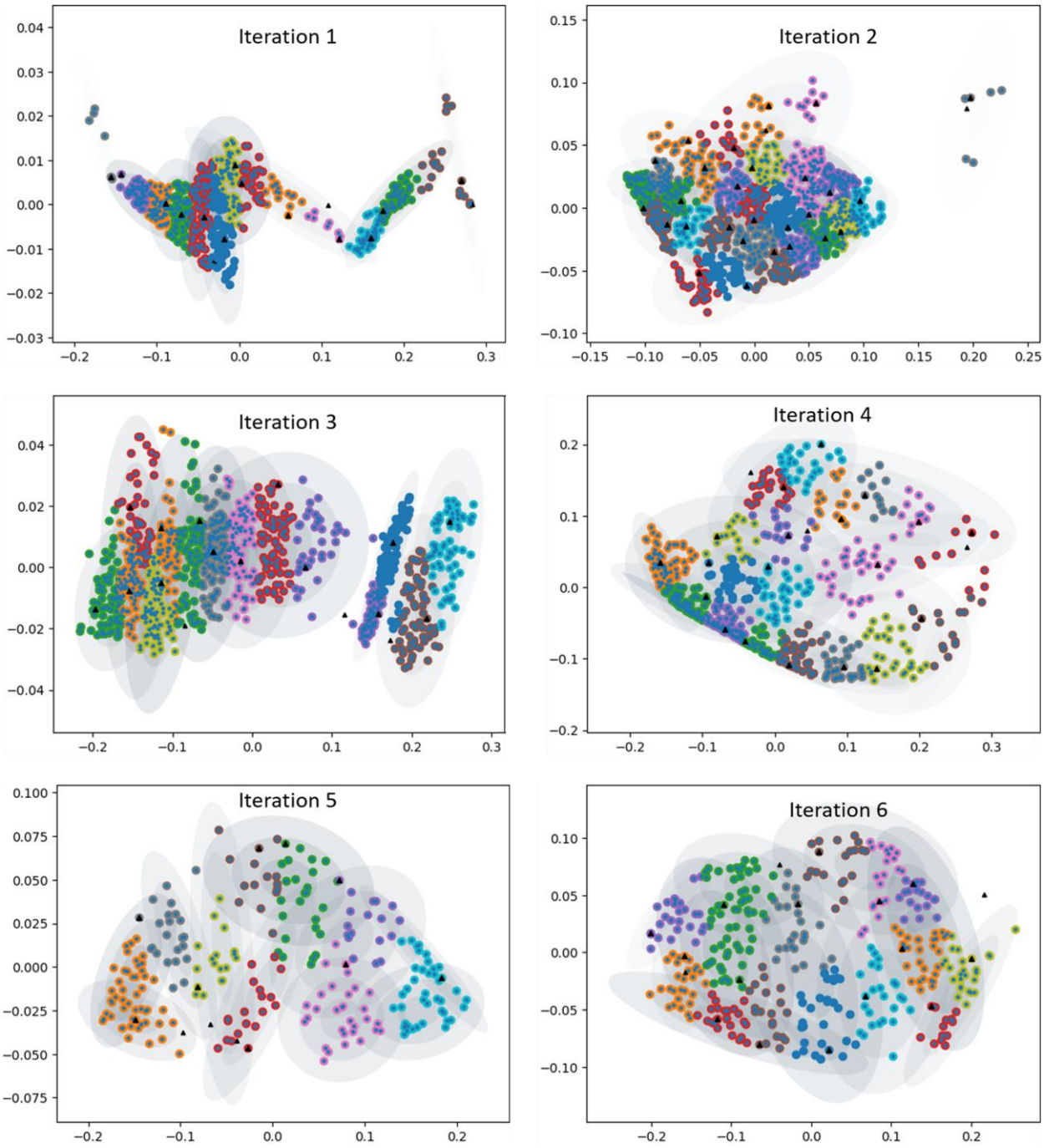

Figure S10. Affinity Propagation on the PCA latent space of the samples. Each color represents a cluster. The black triangles (10-30) are the alloys selected for the second stage of TERM.

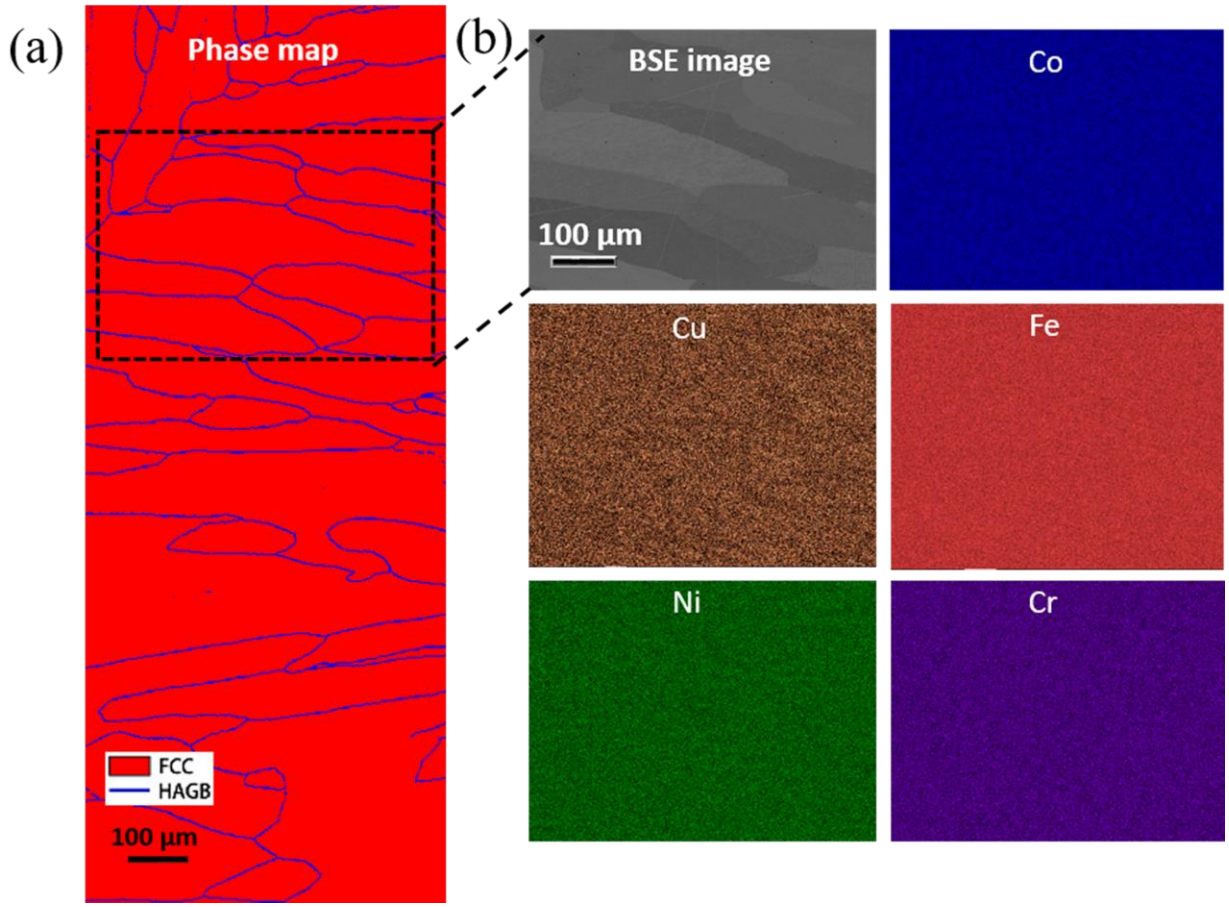

Figure S11. EBSD map, BSE image and corresponding EDS maps of the $Fe_{48.8}Ni_{17.8}Co_{22.2}Cr_{6.2}Cu_5$ (wt.%) HEA (B2 in Table 1). The EBSD maps show the alloys are fully recrystallized with only a single fcc structure. The BSE image and EDS maps show Fe, Co, Cr and Ni are uniformly distributed, while Cu has small inhomogeneity.

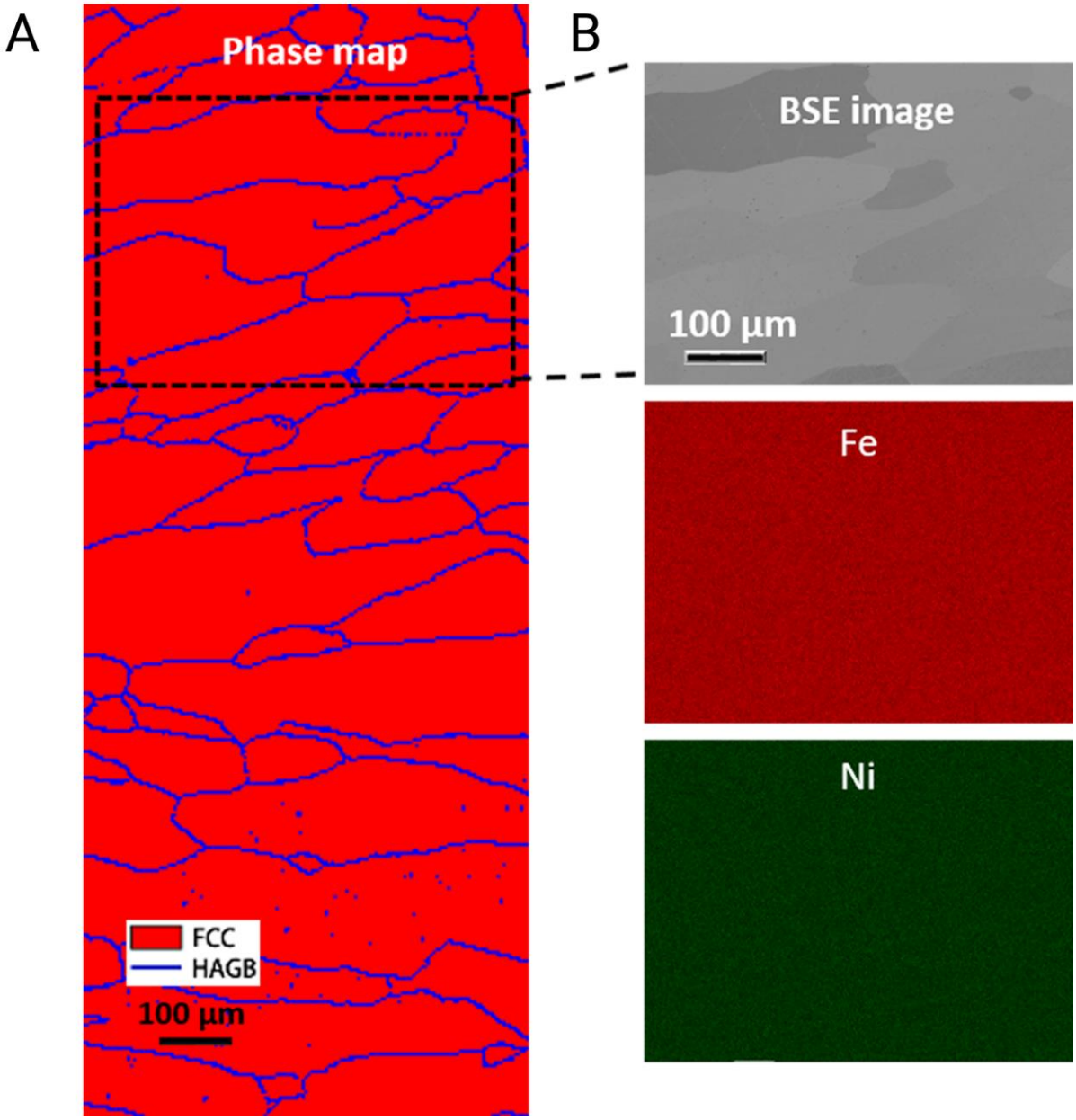

Figure S12. EBSD map, BSE image and corresponding EDS maps of the $Fe_{63.5}Ni_{36.5}$ (wt.%) HEA (A6 in Table 1). The EBSD maps show the alloys are fully recrystallized with only a single fcc structure. The BSE image and EDS maps show Fe and Ni are uniformly distributed.

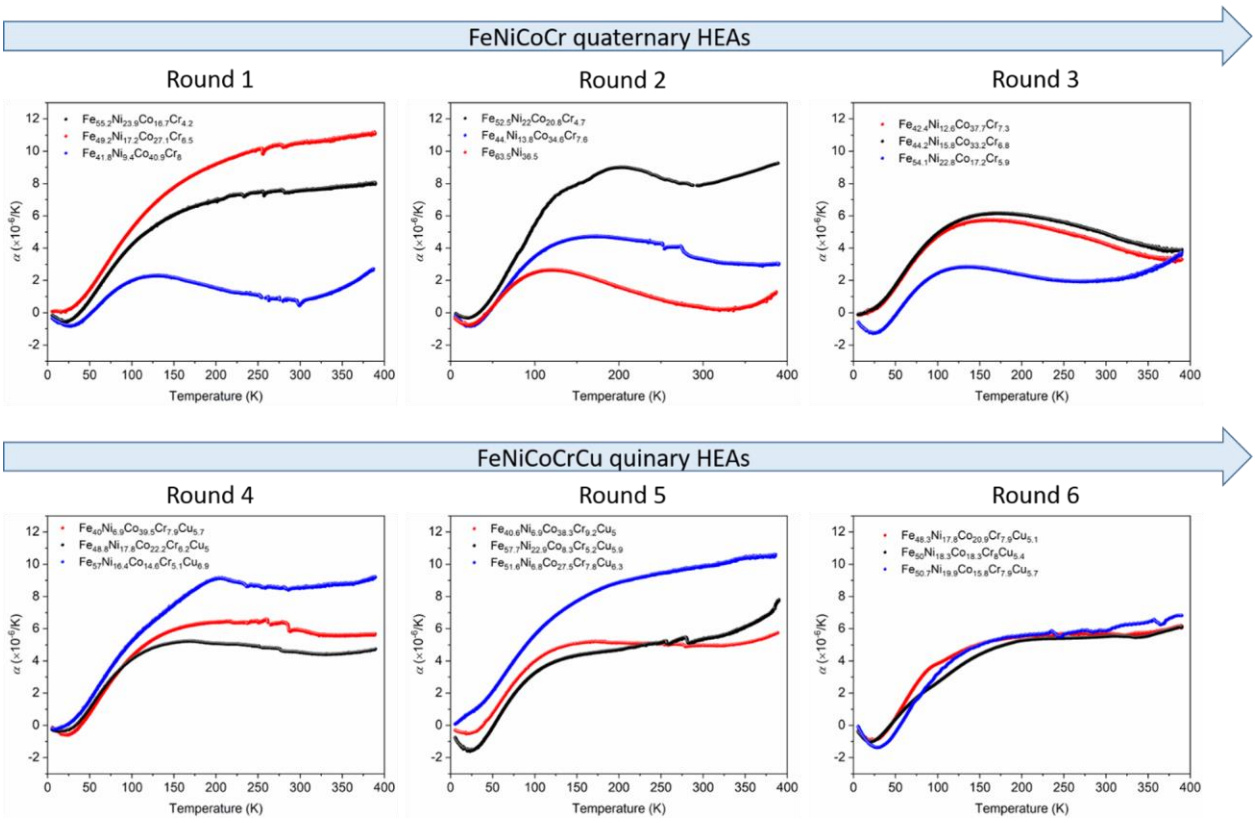

Figure S13. Experimental results of the thermal expansion behavior of the alloys from 1st to 6th rounds.

Table S1. The effect of various treatments on the TEC of $Fe_{64}Ni_{36}$ (wt.%) Invar (*62-63*). The change of TEC (0-0.4×10^{-6} / ºC) of Invar alloy with different heat treatment and rolling processes is relatively small, considering that the gap of TEC between Invar alloy and conventional alloys is more than 10×10^{-6} / ºC.

Treatments					$\Delta\alpha$ (×10^{-6} / ºC) (change of TEC)
Hot-rolled at 1200 ºC					0
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled			0
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Water quenched			0
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled or water quenched	30-98% cold rolled		0.2-0.4
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled or water quenched	30-98% cold rolled	Aged at 100 ºC	0.2-0.4
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled or water quenched	30-98% cold rolled	Aged at 500 ºC	0-0.4
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled or water quenched	30-98% cold rolled	Aged at 500 ºC	0
Hot-rolled at 1200 ºC	Annealed at 1000 ºC	Furnace cooled or water quenched	97% cold rolled	Heated at 1000 ºC	0

			(cube textured)		
Hot-rolled at 1200 °C	Annealed at 1000 °C	Furnace cooled or water quenched	97% cold rolled	Heated at 1000 °C (cube textured)	Cold-rolled	0.7
Hot-rolled at 1200 °C	Annealed at 1000 °C	Water quenched	50% cold rolled	Aged at 100 °C		0.17
Hot-rolled at 1200 °C	Annealed at 1000 °C	Water quenched	95% cold rolled	Aged at 100 °C		0.15

Table S2. Comparison of the efficiency of this work with try and error. Try and error's discovery rate decreases remarkably from binary to ternary alloy system (from 5.3% to 1.1%) *(34-39).* For multi-principal alloys. The composition space grows exponentially in terms of a multi-principal alloy system. Nevertheless, HEA-GAD-TERMS demonstrates a very high discovery rate (25%).

Method	Alloy system	Discovered Invar compositions (TEC≤2×10-6 K^{-1})	Total experimental compositions	Discovery rate (%)
Try and error	Fe-Ni	$Fe_{63.5}Ni_{36.5}$ (classic Invar)	19	5.3
	Fe-Co	0	17	0
	Co-Ni	0	13	0
	Fe-Ni-Co	$Fe_{63.5}Ni_{31.5}Co_5$ (super Invar)	62	1.6
	Fe-Co-Cr	$Fe_{37}Co_{54}Cr_9$ (stainless Invar)	93	1.1
ML-enabled alloy design	***Fe-Ni-Co-Cr***	***$Fe_{41.8}Ni_{9.4}Co_{40.9}Cr_8$ $Fe_{54.1}Ni_{22.8}Co_{17.2}Cr_{5.9}$ (HEA Invar)***	***8***	***25***

Table S3. FeNiCoCrCu HEAs designed with change of VECs in our previous work.

Alloys	Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	Cu (wt. %)	VEC	TEC ($\times 10^{-6}$/K)
1	0.602	0.241	0.093	0.007	0.057	8.70	8.83
2	0.569	0.129	0.154	0.050	0.098	8.55	9.90
3	0.466	0.051	0.351	0.080	0.052	8.40	9.45

Table S4. The first 6 compositions in the dataset showing the structure of the attributes.

Compositions		**Atomic features***					**Thermodynamic calculations**		**DFT calculations**
Fe	Ni	Valence electron concentration (VEC)	Atomic radius (empirical) (pm)	Covalent radius (pm)	Pauling electron egativity (Pauling units)	…	Magnetostriction	Curie temperature (K)	Magnetization (Bohr magneton)
1	0	8	140	124	1.83	…	0.0067	1043	2.22
0.95	0.05	8.0954	139.7616	124.047	1.8338	…	0.0166	1029.065	2.1792
0.9	0.1	8.1912	139.5219	124.095	1.8376	…	0.0251	1015.874	2.1406
0.85	0.15	8.2875	139.2812	124.143	1.8415	…	0.0254	1002.941	2.1027
0.8	0.2	8.3843	139.0392	124.192	1.8454	…	0.0232	989.8446	2.0644
0.75	0.25	8.4816	138.796	124.240	1.8493	…	0.0214	976.2046	2.0245

*****The atomic features include valence electron concentration, atomic radius (empirical), covalent radius, Pauling electronegativity, density, thermal conductivity, melting point, first ionization energy, second ionization energy, third ionization energy and magnetic moment.

Table S5. Top 10 hyperparameters of the Bayesian optimization (BO) results for the neural network in the first stage of TERM.

Number	Target (-error)	Batch_size	Learning rate	Nodes of hidden layers	Number of hidden layers
1	-0.095	52.523	0.00050	173.771	2.000
2	-0.098	51.948	0.00050	209.737	8.437
3	-0.100	48.625	0.00061	146.910	10.000
4	-0.116	64.000	0.00050	129.291	10.000
5	-0.116	50.155	0.00100	218.000	2.000
6	-0.117	32.000	0.00050	113.836	2.000
7	-0.117	41.789	0.00050	127.379	4.527
8	-0.118	32.000	0.00084	186.892	2.000
9	-0.119	32.935	0.00050	217.642	9.430
10	-0.119	51.538	0.00099	40.947	10.000

Table S6. Top 10 hyperparameters of the BO results for gradient boosting decision trees (GBDT) in the first stage of TERM.

Number	Target(-error)	Colsample_bytree	Learning rate	Max_bin	…*
1	-0.120	0.830	0.394	91.466	…
2	-0.121	0.725	0.204	80.118	…
3	-0.126	0.820	0.231	89.541	…
4	-0.129	0.593	0.192	80.770	…
5	-0.130	0.843	0.263	75.054	…
6	-0.131	0.733	0.084	83.871	…
7	-0.133	0.898	0.157	76.317	…
8	-0.137	0.962	0.546	77.773	…
9	-0.137	0.768	0.380	75.639	…
10	-0.138	0.772	0.246	40.801	…

*the total hyperparameters includes: colsample_bytree, learning_rate, max_bin, max_depth, min_child_samples, min_child_weight, min_split_gain, n_estimators, num_leaves, reg_alpha, reg_lambda, subsample

Table S7. Top 10 hyperparameters of the BO results for the NN in the second stage of TERM.

Number	Target	Batch_size	Learning rate	Nodes of hidden layers	Number of hidden layers
1	-0.116	47.056	0.00083	204.441	2.056
2	-0.128	62.927	0.00082	194.886	2.080
3	-0.130	45.585	0.00082	215.188	2.124
4	-0.131	63.560	0.00079	217.075	5.347
5	-0.131	52.632	0.00050	105.568	9.944
6	-0.134	36.696	0.00055	53.625	4.764
7	-0.134	32.549	0.00093	76.439	2.042
8	-0.134	63.904	0.00061	150.592	2.007
9	-0.135	43.117	0.00054	208.298	5.080
10	-0.136	63.838	0.00090	185.680	6.069

Table S8. Top 10 hyperparameters of the BO results for GBDT in the second stage of TERM.

Number	Target(-error)	Colsample_bytree	Learning rate	Max_bin	…*
1	-0.151	0.758	0.478	19.504	…
2	-0.162	0.932	0.323	68.725	…
3	-0.162	0.551	0.415	70.968	…
4	-0.170	0.624	0.263	76.266	…
5	-0.177	0.543	0.797	31.433	…
6	-0.178	1.000	1.000	43.033	…
7	-0.182	0.726	0.376	97.658	…
8	-0.183	1.000	1.000	48.325	…
9	-0.184	0.695	0.487	62.409	…
10	-0.184	1.000	1.000	17.432	…

*the total hyperparameters includes: colsample_bytree, learning_rate, max_bin, max_depth, min_child_samples, min_child_weight, min_split_gain, n_estimators, num_leaves, reg_alpha, reg_lambda, subsample

Table S9. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the first iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean ($\times 10^{-6}$/K) (rank)	Prediction std ($\times 10^{-6}$/K) (rank)	*Total ranking score(α=0.9, β=0.1)
0.492	0.172	0.271	0.065	0	0	3.13 (1)	0.75 (15)	2.4
0.552	0.239	0.167	0.042	0	0	3.41 (2)	1.29 (12)	3.0
0.418	0.094	0.409	0.079	0	0	4.39 (3)	0.79 (14)	4.1
0.538	0.215	0.187	0.059	0	0	5.96 (4)	1.35 (11)	4.7
0.558	0.239	0.155	0.049	0	0	6.03 (5)	2.00 (7)	5.2
0.580	0.257	0.117	0.046	0	0	6.88 (6)	5.74 (1)	5.5
0.542	0.217	0.180	0.061	0	0	7.51 (7)	2.60 (5)	6.8
0.546	0.232	0.175	0.048	0	0	8.71 (8)	1.44 (10)	8.2
0.606	0.276	0.073	0.045	0	0	11.38 (9)	3.96 (2)	8.3
0.612	0.282	0.063	0.043	0	0	11.42 (10)	3.94 (3)	9.3
0.511	0.196	0.233	0.060	0	0	11.69 (11)	0.88 (13)	11.2
0.481	0.173	0.284	0.063	0	0	13.43 (12)	1.61 (9)	11.7
0.458	0.147	0.322	0.073	0	0	13.63 (13)	2.24 (6)	12.3
0.415	0.103	0.394	0.088	0	0	13.95 (14)	2.87 (4)	13.0
0.463	0.156	0.313	0.067	0	0	13.98 (15)	1.96 (8)	14.3

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction

Table S10. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the second iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean (×10^{-6}/K) (rank)	Prediction std (×10^{-6}/K) (rank)	*Total ranking score(α=0.9, β=0.1)
*0.635	0.365	0	0	0	0	1.57 (1)	0.51 (30)	3.9
0.440	0.137	0.347	0.076	0	0	4.20 (3)	1.56 (18)	4.5
0.525	0.219	0.208	0.048	0	0	3.91 (2)	1.03 (28)	4.6
0.556	0.272	0.093	0.079	0	0	7.52 (5)	2.52 (3)	4.8
0.575	0.251	0.114	0.061	0	0	7.53 (6)	2.88 (1)	5.5
0.548	0.234	0.144	0.075	0	0	6.45 (4)	1.21 (25)	6.1
0.542	0.255	0.119	0.084	0	0	7.65 (7)	2.03 (8)	7.1
0.526	0.215	0.175	0.085	0	0	8.19 (8)	1.48 (21)	9.3
0.569	0.231	0.141	0.06	0	0	8.88 (10)	2.39 (4)	9.4
0.523	0.227	0.16	0.089	0	0	8.82 (9)	1.53 (20)	10.1
0.537	0.269	0.104	0.09	0	0	9.19 (11)	1.84 (11)	11.0
0.544	0.215	0.169	0.073	0	0	9.55 (12)	2.10 (6)	11.4
0.527	0.262	0.115	0.096	0	0	10.51 (14)	2.11 (5)	13.1

0.496	0.198	0.206	0.101	0	0	10.15 (13)	1.07 (27)	14.4
0.528	0.167	0.234	0.071	0	0	10.76 (15)	1.55 (19)	15.4
0.549	0.187	0.203	0.062	0	0	10.81 (16)	1.68 (17)	16.1
0.499	0.233	0.161	0.108	0	0	11.13 (17)	1.25 (22)	17.5
0.507	0.245	0.143	0.105	0	0	11.56 (18)	1.69 (16)	17.8
0.458	0.104	0.335	0.102	0	0	11.92 (21)	2.05 (7)	19.6
0.477	0.193	0.218	0.112	0	0	11.63 (19)	0.92 (29)	20.0
0.476	0.209	0.198	0.117	0	0	11.75 (20)	1.23 (23)	20.3
0.494	0.244	0.148	0.114	0	0	12.88 (23)	1.80 (13)	22.0
0.479	0.222	0.181	0.118	0	0	12.16 (22)	1.21 (26)	22.4
0.498	0.273	0.111	0.118	0	0	13.50 (24)	1.83 (12)	22.8
0.459	0.184	0.234	0.122	0	0	13.82 (25)	1.23 (24)	24.9
0.473	0.236	0.165	0.126	0	0	14.13 (26)	1.74 (15)	24.9
0.47	0.259	0.137	0.134	0	0	15.17 (28)	2.61 (2)	25.4
0.469	0.228	0.176	0.127	0	0	14.42 (27)	1.79 (14)	25.7
0.453	0.203	0.214	0.131	0	0	15.85 (29)	1.87 (10)	27.1
0.456	0.211	0.201	0.131	0	0	16.12 (30)	1.96(9)	27.9

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction.

*The first alloy is not sampled from the model, but the original FeNi Invar as a reference alloy.

Table S11. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the third iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean ($\times 10^{-6}$/K) (rank)	Prediction std ($\times 10^{-6}$/K) (rank)	*Total ranking score(α=0.9, β=0.1)
0.424	0.126	0.377	0.073	0	0	4.58 (1)	1.40 (14)	2.3
0.541	0.228	0.173	0.059	0	0	5.16 (2)	1.43 (13)	3.1
0.443	0.158	0.332	0.068	0	0	5.88 (3)	2.17 (9)	3.6
0.549	0.238	0.142	0.071	0	0	6.98 (4)	6.25 (1)	4.6
0.512	0.214	0.19	0.083	0	0	6.17 (5)	1.25 (15)	5.1
0.569	0.254	0.111	0.066	0	0	8.01 (6)	5.42 (2)	5.6
0.499	0.199	0.218	0.084	0	0	8.89 (7)	1.43 (12)	7.5
0.514	0.225	0.174	0.088	0	0	9.41 (8)	1.89 (11)	8.3
0.537	0.244	0.137	0.082	0	0	10.16 (9)	3.90 (4)	8.5
0.592	0.269	0.08	0.059	0	0	10.18 (10)	4.17 (3)	9.3
0.55	0.262	0.104	0.083	0	0	11.81 (11)	3.19 (5)	11.3
0.461	0.188	0.247	0.104	0	0	11.54 (12)	1.23 (16)	11.5
0.408	0.121	0.366	0.105	0	0	12.40 (13)	2.60 (8)	13.4
0.463	0.164	0.282	0.09	0	0	11.99 (14)	0.84 (17)	13.4
0.432	0.121	0.358	0.089	0	0	12.89 (15)	2.60 (7)	14.2

0.405	0.096	0.406	0.094	0	0	13.18 (16)	3.07 (6)	15.0
0.371	0.094	0.419	0.116	0	0	14.44 (17)	2.16 (10)	16.3

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction.

Table S12. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the fourth iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean ($\times 10^{-6}$/K) (rank)	Prediction std ($\times 10^{-6}$/K) (rank)	*Total ranking score(α=0.9, β=0.1)
0.570	0.164	0.146	0.051	0	0.069	4.43 (1)	1.33 (17)	2.6
0.488	0.178	0.222	0.063	0	0.050	5.48 (2)	1.01 (19)	3.7
0.400	0.069	0.395	0.079	0	0.058	7.57 (3)	1.45 (16)	4.3
0.490	0.182	0.191	0.089	0	0.049	7.70 (4)	1.50 (13)	4.9
0.498	0.261	0.109	0.100	0	0.032	8.95 (5)	2.64 (3)	5.7
0.468	0.221	0.175	0.102	0	0.034	8.87 (6)	1.27 (18)	6.3
0.520	0.296	0.057	0.099	0	0.028	9.07 (7)	2.60 (4)	6.7
0.426	0.302	0.138	0.128	0	0.006	9.46 (8)	1.50 (14)	8.6
0.467	0.289	0.112	0.114	0	0.018	9.97 (9)	2.22 (9)	9.0

0.406	0.297	0.161	0.134	0	0.003	10.20 (10)	1.46 (15)	10.5
0.539	0.146	0.179	0.068	0	0.068	11.22 (11)	2.26 (8)	10.7
0.495	0.093	0.269	0.072	0	0.071	12.14 (12)	2.06 (10)	11.8
0.564	0.085	0.212	0.050	0	0.088	12.57 (13)	2.33 (5)	12.2
0.482	0.129	0.247	0.082	0	0.059	12.67 (14)	2.29 (7)	13.3
0.422	0.139	0.293	0.102	0	0.043	13.00 (15)	2.31 (6)	14.1
0.411	0.046	0.391	0.089	0	0.063	13.10 (16)	2.85 (1)	14.5
0.445	0.075	0.332	0.084	0	0.064	13.34 (17)	2.72 (2)	15.5
0.321	0.223	0.309	0.146	0	0.001	14.17 (18)	1.69 (12)	17.4
0.349	0.151	0.350	0.126	0	0.024	14.95 (19)	1.99 (11)	18.2

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction.

Table S13. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the fifth iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean ($\times 10^{-6}$/K) (rank)	Prediction std ($\times 10^{-6}$/K) (rank)	*Total ranking score(α=0.9, β=0.1)
0.577	0.229	0.083	0.052	0	0.059	4.50 (1)	1.00 (17)	2.6
0.516	0.068	0.275	0.078	0	0.063	9.32 (3)	3.49 (2)	3.7
0.406	0.069	0.383	0.091	0	0.050	8.41 (2)	1.70 (12)	4.3
0.484	0.202	0.162	0.069	0	0.083	9.33 (4)	1.81 (11)	4.9
0.523	0.174	0.142	0.055	0	0.106	9.51 (5)	3.46 (3)	5.7
0.503	0.253	0.093	0.068	0	0.083	9.58 (6)	2.35 (6)	6.3
0.473	0.109	0.261	0.064	0	0.093	9.94 (7)	2.65 (4)	6.7
0.446	0.203	0.205	0.081	0	0.064	10.14 (8)	1.25 (16)	8.6
0.464	0.252	0.140	0.080	0	0.064	10.55 (9)	1.54 (14)	9.0
0.407	0.029	0.413	0.077	0	0.074	11.80 (10)	2.10 (7)	10.5
0.447	0.062	0.335	0.068	0	0.088	11.89 (11)	1.87 (10)	10.7
0.405	0.208	0.250	0.094	0	0.043	11.99 (12)	1.64 (13)	11.8
0.412	0.112	0.330	0.083	0	0.063	12.17 (13)	2.10 (8)	12.2
0.384	0.159	0.319	0.096	0	0.041	12.60 (15)	2.60 (5)	13.3
0.406	0.063	0.382	0.080	0	0.068	12.30 (14)	1.28 (15)	14.1

0.508	0.095	0.233	0.052	0	0.112	12.88 (16)	4.23 (1)	14.5
0.384	0.119	0.358	0.093	0	0.048	13.82 (17)	1.87 (9)	15.5

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction.

Table S14. Sampled compositions for the physically-informed model of the Two-stage Ensemble Regression Model (TERM). Candidates in the sixth iteration.

Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)	V (wt. %)	Cu (wt. %)	prediction mean ($\times10^{-6}$/K) (rank)	Prediction std ($\times10^{-6}$/K) (rank)	*Total ranking score(α=0.9, β=0.1)
0.483	0.178	0.209	0.079	0	0.051	5.49 (1)	0.92 (12)	2.1
0.507	0.199	0.158	0.079	0	0.057	5.56 (2)	1.03 (11)	2.9
0.500	0.183	0.183	0.080	0	0.054	5.65 (3)	1.16 (10)	3.7
0.574	0.177	0.126	0.049	0	0.072	8.90 (4)	3.35 (3)	3.9
0.523	0.178	0.178	0.060	0	0.061	9.33 (5)	1.18 (8)	5.3
0.539	0.222	0.128	0.053	0	0.058	9.60 (6)	1.82 (7)	6.1
0.463	0.096	0.302	0.078	0	0.061	12.06 (7)	1.18 (9)	7.2
0.414	0.054	0.383	0.092	0	0.056	12.93 (8)	2.47 (4)	7.6
0.478	0.179	0.223	0.069	0	0.051	13.30 (9)	2.27 (6)	8.7

0.496	0.064	0.293	0.074	0	0.073	13.91 (10)	2.44 (5)	9.5
0.544	0.090	0.225	0.062	0	0.079	16.37 (11)	4.07 (2)	10.1
0.534	0.075	0.247	0.065	0	0.079	17.06 (12)	4.29 (1)	10.9

*Score = α×Rank in TEC prediction+ β× Rank in Uncertainty prediction.

Table S15. Thermodynamic and DFT calculations of the candidates in the 1-6 iterations.

Iteration1	Fe	Ni	Co	Cr	V	Cu	Phase	Curie temperature (Thermodynamic calculation) (K)	Magnetic moment (Thermodynamic calculation) (μ_B/atom)	Magnetostriction (DFT calculation)
1	0.552	0.239	0.167	0.042	0	0	'FCC_A1'	412.29	1.74	0.043
2	0.492	0.172	0.271	0.065	0	0	'FCC_A1'	400.97	1.70	0.046
3	0.418	0.094	0.409	0.079	0	0	'FCC_A1'	467.12	1.66	0.050
4	0.580	0.257	0.117	0.046	0	0	'FCC_A1'	366.12	1.58	0.041

5	0.557	0.239	0.155	0.049	0	0	'FCC_A1'	379.32	1.64	0.042
6	0.606	0.276	0.073	0.045	0	0	'FCC_A1'	345.24	1.45	0.041
7	0.415	0.103	0.394	0.088	0	0	'FCC_A1'	437.17	1.59	0.048
8	0.480	0.173	0.284	0.063	0	0	'FCC_A1'	426.65	1.74	0.045
9	0.612	0.282	0.063	0.043	0	0	'FCC_A1'	348.32	1.44	0.040
10	0.542	0.217	0.180	0.061	0	0	'FCC_A1'	349.01	1.58	0.043
11	0.464	0.156	0.313	0.067	0	0	'FCC_A1'	435.72	1.73	0.046
12	0.545	0.232	0.175	0.048	0	0	'FCC_A1'	397.76	1.70	0.042
13	0.511	0.196	0.233	0.060	0	0	'FCC_A1'	393.61	1.69	0.044
14	0.458	0.147	0.322	0.073	0	0	'FCC_A1'	422.06	1.68	0.046
15	0.539	0.215	0.187	0.059	0	0	'FCC_A1'	360.38	1.62	0.043

Iteration2										
1	0.525	0.219	0.2083	0.0475	0	0	'FCC_A1'	424.59	1.77	0.044
2	0.440	0.137	0.3467	0.0758	0	0	'FCC_A1'	439.23	1.68	0.047
3	0.635	0.365	0	0	0	0	'FCC_A1'	535.74	1.64	0.037
4	0.547	0.234	0.144	0.075	0	0	'FCC_A1'	292.82	1.39	0.041
5	0.556	0.272	0.093	0.079	0	0	'FCC_A1'	274.75	1.25	0.038
6	0.498	0.233	0.161	0.108	0	0	'FCC_A1'	230.78	1.19	0.038
7	0.524	0.227	0.160	0.089	0	0	'FCC_A1'	268.84	1.32	0.040
8	0.537	0.269	0.104	0.090	0	0	'FCC_A1'	254.79	1.20	0.037
9	0.476	0.209	0.198	0.117	0	0	'FCC_A1'	225.42	1.19	0.039
10	0.574	0.251	0.114	0.061	0	0	'FCC_A1'	315.22	1.43	0.041

11	0.525	0.215	0.175	0.085	0	0	'FCC_A1'	281.27	1.38	0.042
12	0.527	0.262	0.115	0.096	0	0	'FCC_A1'	242.91	1.18	0.037
13	0.479	0.222	0.181	0.118	0	0	'FCC_A1'	217.15	1.16	0.038
14	0.548	0.187	0.203	0.062	0	0	'FCC_A1'	331.84	1.58	0.046
15	0.542	0.255	0.119	0.084	0	0	'FCC_A1'	269.08	1.28	0.039
16	0.498	0.273	0.111	0.118	0	0	'FCC_A1'	194.17	1.02	0.034
17	0.477	0.193	0.218	0.112	0	0	'FCC_A1'	243.74	1.26	0.040
18	0.507	0.245	0.143	0.105	0	0	'FCC_A1'	231.43	1.18	0.037
19	0.495	0.198	0.206	0.101	0	0	'FCC_A1'	260.56	1.32	0.041
20	0.494	0.244	0.148	0.114	0	0	'FCC_A1'	213.33	1.12	0.037
21	0.452	0.203	0.214	0.131	0	0	'FCC_A1'	204.44	1.12	0.037

22	0.459	0.185	0.234	0.122	0	0	'FCC_A1'	231.21	1.21	0.040
23	0.470	0.259	0.137	0.134	0	0	'FCC_A1'	165.77	0.96	0.033
24	0.568	0.231	0.141	0.060	0	0	'FCC_A1'	321.19	1.50	0.043
25	0.457	0.211	0.201	0.131	0	0	'FCC_A1'	197.15	1.10	0.037
26	0.543	0.215	0.169	0.073	0	0	'FCC_A1'	302.26	1.45	0.043
27	0.469	0.228	0.176	0.127	0	0	'FCC_A1'	195.50	1.09	0.036
28	0.473	0.236	0.165	0.126	0	0	'FCC_A1'	193.80	1.07	0.036
29	0.528	0.167	0.234	0.071	0	0	'FCC_A1'	324.74	1.55	0.047
30	0.459	0.104	0.335	0.102	0	0	'FCC_A1'	316.21	1.41	0.048
Iteration3										
1	0.424	0.126	0.377	0.073	0	0	'FCC_A1'	477.42	1.72	0.049

2	0.443	0.158	0.332	0.068	0	0	'FCC_A1'	258.65	0.47	0.047
3	0.541	0.228	0.173	0.059	0	0	'FCC_A1'	604.82	0.48	0.044
4	0.537	0.244	0.137	0.082	0	0	'FCC_A1'	280.82	1.33	0.040
5	0.592	0.269	0.080	0.059	0	0	'FCC_A1'	306.76	1.36	0.040
6	0.551	0.262	0.104	0.083	0	0	'FCC_A1'	263.52	1.24	0.039
7	0.432	0.121	0.358	0.089	0	0	'FCC_A1'	403.41	1.57	0.047
8	0.404	0.096	0.406	0.094	0	0	'FCC_A1'	432.82	1.55	0.048
9	0.499	0.199	0.218	0.084	0	0	'FCC_A1'	318.74	1.47	0.042
10	0.513	0.214	0.190	0.083	0	0	'FCC_A1'	304.40	1.43	0.041
11	0.549	0.238	0.142	0.071	0	0	'FCC_A1'	305.97	1.43	0.041
12	0.371	0.094	0.419	0.116	0	0	'FCC_A1'	399.15	1.40	0.046

13	0.408	0.121	0.366	0.105	0	0	'FCC_A1'	378.31	1.46	0.045
14	0.569	0.254	0.111	0.066	0	0	'FCC_A1'	303.50	1.39	0.041
15	0.513	0.225	0.174	0.088	0	0	'FCC_A1'	286.36	1.37	0.040
16	0.461	0.188	0.247	0.104	0	0	'FCC_A1'	295.17	1.37	0.041
17	0.464	0.164	0.282	0.090	0	0	'FCC_A1'	346.26	1.51	0.044
Iteration4										
1	0.400	0.069	0.395	0.079	0	0.0577	'FCC_A1'	14.49	0.00	0.044
2	0.488	0.178	0.2221	0.0626	0	0.0502	'FCC_A1'	375.39	1.50	0.039
3	0.570	0.164	0.146	0.051	0	0.069	'FCC_A1'	282.98	1.26	0.042
4	0.520	0.296	0.057	0.099	0	0.028	'FCC_A1'	231.11	0.97	0.032
5	0.467	0.289	0.112	0.114	0	0.018	'FCC_A1'	232.75	1.04	0.031

6	0.411	0.046	0.391	0.089	0	0.063	'FCC_A1'	386.29	1.27	0.044
7	0.426	0.302	0.138	0.128	0	0.006	'FCC_A1'	226.44	1.02	0.029
8	0.349	0.151	0.350	0.126	0	0.024	'FCC_A1'	350.04	1.25	0.037
9	0.445	0.075	0.332	0.084	0	0.064	'FCC_A1'	344.20	1.28	0.043
10	0.495	0.093	0.269	0.072	0	0.071	'FCC_A1'	300.40	1.25	0.042
11	0.520	0.234	0.114	0.089	0	0.043	'FCC_A1'	248.53	1.13	0.035
12	0.565	0.085	0.212	0.050	0	0.088	'FCC_A1'	246.37	1.10	0.043
13	0.498	0.261	0.109	0.100	0	0.032	'FCC_A1'	243.32	1.09	0.033
14	0.483	0.129	0.247	0.082	0	0.059	'FCC_A1'	297.37	1.28	0.040
15	0.405	0.297	0.161	0.134	0	0.003	'FCC_A1'	227.02	1.02	0.029
16	0.539	0.146	0.179	0.068	0	0.068	'FCC_A1'	260.99	1.20	0.040

17	0.490	0.181	0.191	0.089	0	0.049	'FCC_A1'	274.41	1.24	0.037
18	0.423	0.139	0.293	0.102	0	0.043	'FCC_A1'	325.19	1.30	0.039
19	0.321	0.223	0.309	0.146	0	0.001	'FCC_A1'	296.82	1.13	0.032
20	0.468	0.221	0.175	0.102	0	0.034	'FCC_A1'	267.25	1.20	0.035
Iteration5										
1	0.406	0.069	0.383	0.0914	0	0.05	'FCC_A1'	12.40	0.00	0.045
2	0.577	0.229	0.083	0.052	0	0.059	'FCC_A1'	305.14	1.25	0.038
3	0.516	0.068	0.275	0.078	0	0.063	'FCC_A1'	2.24	0.01	0.048
4	0.407	0.063	0.382	0.080	0	0.068	'FCC_A1'	418.23	1.36	0.042
5	0.412	0.112	0.330	0.083	0	0.063	'FCC_A1'	397.65	1.39	0.039
6	0.382	0.119	0.358	0.093	0	0.048	'FCC_A1'	418.29	1.41	0.039

7	0.407	0.029	0.413	0.077	0	0.074	'FCC_A1'	427.96	1.30	0.045
8	0.385	0.159	0.319	0.096	0	0.041	'FCC_A1'	400.52	1.41	0.037
9	0.503	0.161	0.178	0.060	0	0.098	'FCC_A1'	318.47	1.26	0.034
10	0.473	0.109	0.261	0.064	0	0.093	'FCC_A1'	343.40	1.30	0.038
11	0.503	0.253	0.093	0.068	0	0.083	'FCC_A1'	325.60	1.19	0.029
12	0.447	0.203	0.205	0.081	0	0.064	'FCC_A1'	350.61	1.34	0.033
13	0.405	0.208	0.250	0.094	0	0.043	'FCC_A1'	371.62	1.37	0.034
14	0.484	0.202	0.162	0.069	0	0.083	'FCC_A1'	332.65	1.29	0.032
15	0.447	0.062	0.335	0.068	0	0.088	'FCC_A1'	373.86	1.29	0.041
16	0.464	0.252	0.140	0.080	0	0.064	'FCC_A1'	336.95	1.27	0.030
17	0.523	0.174	0.142	0.055	0	0.106	'FCC_A1'	306.29	1.20	0.033

18	0.508	0.095	0.233	0.052	0	0.112	'FCC_A1'	308.79	1.19	0.038
Iteration6										
1	0.483	0.178	0.209	0.079	0	0.051	'FCC_A1'	319.12	1.35	0.039
2	0.500	0.183	0.183	0.08	0	0.054	'FCC_A1'	290.47	1.27	0.039
3	0.507	0.199	0.158	0.079	0	0.057	'FCC_A1'	285.21	1.24	0.037
4	0.539	0.222	0.128	0.053	0	0.058	'FCC_A1'	346.34	1.39	0.035
5	0.534	0.075	0.247	0.065	0	0.079	'FCC_A1'	251.28	1.11	0.044
6	0.478	0.179	0.223	0.069	0	0.051	'FCC_A1'	364.03	1.46	0.037
7	0.496	0.064	0.293	0.074	0	0.073	'FCC_A1'	286.91	1.17	0.044
8	0.544	0.090	0.225	0.062	0	0.079	'FCC_A1'	249.26	1.12	0.043
9	0.463	0.096	0.302	0.078	0	0.061	'FCC_A1'	339.73	1.33	0.042

10	0.576	0.177	0.126	0.049	0	0.072	'FCC_A1'	285.12	1.25	0.038
11	0.523	0.178	0.178	0.060	0	0.061	'FCC_A1'	326.49	1.38	0.038
12	0.415	0.054	0.383	0.092	0	0.056	'FCC_A1'	374.85	1.28	0.044

Table S16. Bulk chemical compositions of the alloys in iteration 1 measured by inductively coupled plasma mass spectrometry.

Alloys (wt. %)	Fe (wt. %)	Ni (wt. %)	Co (wt. %)	Cr (wt. %)
$Fe_{55.2}Ni_{23.9}Co_{16.7}Cr_{4.2}$	56.23	23.10	16.00	4.67
$Fe_{49.2}Ni_{17.2}Co_{27.1}Cr_{6.5}$	49.80	16.60	25.80	7.80
$Fe_{41.8}Ni_{9.4}Co_{40.9}Cr_{8}$	43.41	9.03	39.00	8.56

References and Notes

1. J. W. Yeh, S. K. Chen, S. J. Lin, J. Y. Gan, T. S. Chin, T. T. Shun, C. H. Tsau, S. Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Adv. Eng. Mater.* **6**, 299–303 (2004). doi:10.1002/adem.200300567

2. B. Cantor, I. Chang, P. Knight, A. Vincent, Microstructural development in equiatomic multicomponent alloys. *Mater. Sci. Eng. A* **375–377**, 213–218 (2004). doi:10.1016/j.msea.2003.10.257

3. E. P. George, D. Raabe, R. O. Ritchie, High-entropy alloys. *Nat. Rev. Mater.* **4**, 515–534 (2019). doi:10.1038/s41578-019-0121-4

4. H. Mao, H.-L. Chen, Q. Chen, TCHEA1: A thermodynamic database not limited for "high entropy" alloys. *J. Phase Equilibria Diffus.* **38**, 353–368 (2017). doi:10.1007/s11669-017-0570-7

5. F. Körmann, D. Ma, D. D. Belyea, M. S. Lucas, C. W. Miller, B. Grabowski, M. H. Sluiter, "Treasure maps" for magnetic high-entropy-alloys from theory and experiment. *Appl. Phys. Lett.* **107**, 142404 (2015). doi:10.1063/1.4932571

6. Z. Rao, D. Ponge, F. Körmann, Y. Ikeda, O. Schneeweiss, M. Friák, J. Neugebauer, D. Raabe, Z. Li, Invar effects in FeNiCo medium entropy alloys: From an Invar treasure map to alloy design. *Intermetallics* **111**, 106520 (2019). doi:10.1016/j.intermet.2019.106520

7. S. Huang, E. Holmström, O. Eriksson, L. Vitos, Mapping the magnetic transition temperatures for medium-and high-entropy alloys. *Intermetallics* **95**, 80–84 (2018). doi:10.1016/j.intermet.2018.01.016

8. Z. Rao, H. Springer, D. Ponge, Z. Li, Combinatorial development of multicomponent invar alloys via rapid alloy prototyping. *Materialia (Oxf.)* **21**, 101326 (2022). doi:10.1016/j.mtla.2022.101326

9. B. O. Mukhamedov, K. V. Karavaev, I. A. Abrikosov, Machine learning prediction of thermodynamic and mechanical properties of multicomponent Fe-Cr-based alloys. *Phys. Rev. Mater.* **5**, 104407 (2021). doi:10.1103/PhysRevMaterials.5.104407

10. J. Schmidt, L. Pettersson, C. Verdozzi, S. Botti, M. A. L. Marques, Crystal graph attention networks for the prediction of stable materials. *Sci. Adv.* **7**, eabi7948 (2021). doi:10.1126/sciadv.abi7948 Medline

11. Z. Pei, K. A. Rozman, Ö. N. Doğan, Y. Wen, N. Gao, E. A. Holm, J. A. Hawk, D. E. Alman, M. C. Gao, Machine-learning microstructure for inverse material design. *Adv. Sci. (Weinh.)* **8**, e2101207 (2021). doi:10.1002/advs.202101207 Medline

12. P. V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat. Commun.* **9**, 1668 (2018). doi:10.1038/s41467-018-03821-9 Medline

13. J. Gubernatis, T. Lookman, Machine learning in materials design and discovery: Examples from the present and suggestions for the future. *Phys. Rev. Mater.* **2**, 120301 (2018). doi:10.1103/PhysRevMaterials.2.120301

14. D. Xue, P. V. Balachandran, J. Hogden, J. Theiler, D. Xue, T. Lookman, Accelerated search for materials with targeted properties by adaptive design. *Nat. Commun.* **7**, 11241 (2016). doi:10.1038/ncomms11241 Medline

15. R. Yuan, Z. Liu, P. V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue, T. Lookman, Accelerated discovery of large electrostrains in $BaTiO_3$ based piezoelectrics using active learning. *Adv. Mater.* **30**, 1702884 (2018). doi:10.1002/adma.201702884 Medline

16. E. Wassermann, M. Acet, "Invar and Anti-Invar: Magnetovolume effects in Fe-based alloys revisited," in *Magnetism and Structure in Functional Materials*, A. Planes, L. Mañosa, A. Saxena, Eds. (Springer, 2005), pp. 177–197.

17. M. Shiga, Invar alloys. *Curr. Opin. Solid State Mater. Sci.* **1**, 340–348 (1996). doi:10.1016/S1359-0286(96)80023-4

18. C. L. Lin, T. W. Chen, Y. J. Chang, H. Murakami, S. Mitani, A. C. Yeh, Dimensional stability of a metastable FCC high entropy alloy. *Appl. Phys. Lett.* **119**, 171902 (2021). doi:10.1063/5.0064544

19. Z. Rao, A. Çakır, Ö. Özgün, D. Ponge, D. Raabe, Z. Li, M. Acet, 3*d* transition-metal high-entropy Invar alloy developed by adjusting the valence-electron concentration. *Phys. Rev. Mater.* **5**, 044406 (2021). doi:10.1103/PhysRevMaterials.5.044406

20. C.-L. Lin, J.-L. Lee, S.-M. Kuo, M.-Y. Li, G. Lu, H. Murakami, S. Mitani, S. Gorsse, A.-C. Yeh, Investigation on the thermal expansion behavior of FeCoNi and $Fe_{30}Co_{30}Ni_{30}Cr_{10-x}Mn_x$ high entropy alloys. *Mater. Chem. Phys.* **271**, 124907 (2021). doi:10.1016/j.matchemphys.2021.124907

21. I. Tolstikhin, O. Bousquet, S. Gelly, B. Schoelkopf, Wasserstein auto-encoders. arXiv:1711.01558 [stat.ML] (2017).

22. R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, A. Aspuru-Guzik, Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **4**, 268–276 (2018). doi:10.1021/acscentsci.7b00572 Medline

23. A. Christophe, N. de Freitas, A. Doucet, M. I. Jordan, An introduction to MCMC for machine learning. *Mach. Learn.* **50**, 5–43 (2003). doi:10.1023/A:1020281327116

24. W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109 (1970). doi:10.1093/biomet/57.1.97

25. K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989). doi:10.1016/0893-6080(89)90020-8

26. R. Kohavi, D. H. Wolpert, Bias plus variance decomposition for zero-one loss functions. *ICML* **96**, 275–283 (1996).

27. U. Von Luxburg, B. Schölkopf, "Statistical learning theory: Models, concepts, and results," in *Handbook of the History of Logic*, D. M. Gabbay, S. Hartmann, J. Woods, Eds. (Elsevier, 2011), vol. 10, pp. 651–706.

28. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T. Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, December 2017*, U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, Eds. (NIPS, 2017), p. 3149–3157.

29. J. H. Friedman, Stochastic gradient boosting. *Comput. Stat. Data Anal.* **38**, 367–378 (2002). doi:10.1016/S0167-9473(01)00065-2

30. A. H. Li, J. Bradic, Boosting in the presence of outliers: Adaptive classification with nonconvex loss functions. *J. Am. Stat. Assoc.* **113**, 660–674 (2018). doi:10.1080/01621459.2016.1273116

31. Materials and methods are available as supplementary materials.

32. F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in Statistics*, S. Kotz, N. L. Johnson, Eds. (Springer, 1992), pp. 196–202.

33. R. Lowry, "Concepts and applications of inferential statistics" (DOER, 2014); http://doer.col.org/handle//123456789/4853.

34. H. Masumoto, On the thermal expansion of the alloys of iron, nickel, and cobalt and the cause of the small expansibility of alloys of the invar type. *Science Reports of the Tohoku Imperial University* **20**, 101–123 (1931).

35. H. Masumoto, On the thermal expansion of alloys of cobalt iron and chromium and a new alloy'stainless invar'. *Science Reports of the Tohoku Imperial University* **23**, 265–275 (1934).

36. M. Hakaru, S. Hideo, K. Tatsuo, Influence of addition of nickel on the thermal expansion, rigidity modulus and its temperature coefficient of the alloys of cobalt, iron and chromium, especially of co-elinvar. I: Additions of 10 and 20 per cent of nickel. *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **6**, 529–538 (1954). doi:10.50974/00041753

37. M. Hakaru, S. Hideo, S. Yutaka, Influence of addition of nickel on the thermal expansion, rigidity modulus and its temperature coefficient of the alloys of cobalt, iron and chromium, especially of co-elinvar. II: Additions of 30 and 40 percent of nickel. *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **7**, 533–540 (1955). doi:10.50974/00041822

38. M. Hakaru, S. Hideo, G. Kimiyosi, Influence of addition of nickel on the thermal expansion, the rigidity modulus and its temperature coefficient of the alloys of cobalt, iron and vanadium. *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **9**, 159–169 (1957). doi:10.50974/00041905

39. M. Hakaru, S. Hideo, S. Yutaka, Influence of addition of copper on the characteristics of an elinvar type alloy "co-elinvar." I: Alloys containing 5 percent of copper. *Science Reports of the Research Institutes, Tohoku University. Series A, Physics, Chemistry and Metallurgy* **9**, 170–175 (1957). doi:10.50974/000419066

40. I. A. Abrikosov, A. E. Kissavos, F. Liot, B. Alling, S. Simak, O. Peil, A. V. Ruban, Competition between magnetic structures in the Fe rich fcc FeNi alloys. *Phys. Rev. B Condens. Matter Mater. Phys.* **76**, 014434 (2007). doi:10.1103/PhysRevB.76.014434

41. A. V. Ruban, First-principles modeling of the Invar effect in $Fe_{65}Ni_{35}$ by the spin-wave method. *Phys. Rev. B* **95**, 174432 (2017). doi:10.1103/PhysRevB.95.174432

42. G. Laplanche, P. Gadaud, C. Bärsch, K. Demtröder, C. Reinhart, J. Schreuer, E. George, Elastic moduli and thermal expansion coefficients of medium-entropy subsystems of the CrMnFeCoNi high-entropy alloy. *J. Alloys Compd.* **746**, 244–255 (2018). doi:10.1016/j.jallcom.2018.02.251

43. T. Schneider, M. Acet, B. Rellinghaus, E. F. Wassermann, W. Pepperhoff, Antiferromagnetic Invar and anti-Invar in Fe-Mn alloys. *Phys. Rev. B Condens. Matter* **51**, 8917–8921 (1995). doi:10.1103/PhysRevB.51.8917 Medline

44. C. Chanmuang, M. Naksata, T. Chairuangsri, H. Jain, C. E. Lyman, Microscopy and strength of borosilicate glass-to-Kovar alloy joints. *Mater. Sci. Eng. A* **474**, 218–224 (2008). doi:10.1016/j.msea.2007.04.016

45. K. Fukamichi, M. Kikuchi, S. Arakawa, T. Masumoto, Invar-type new ferromagnetic amorphous Fe-B alloys. *Solid State Commun.* **23**, 955–958 (1977). doi:10.1016/0038-1098(77)90724-4

46. Code for: Z. Rao, P.-Y. Tung, R. Xie, Y. Wei, H. Zhang, A. Ferrari, T. P. C. Klaver, Machine learning-enabled high-entropy alloy discovery, GitHub (2022); https://github.com/ziyuanrao11/Machine-learning-enabled-high-entropy-alloy-discovery.

47. Data for: Z. Rao, P.-Y. Tung, R. Xie, Y. Wei, H. Zhang, A. Ferrari, T. P. C. Klaver, Machine learning-enabled high-entropy alloy discovery, Zenodo (2022); https://doi.org/10.5281/zenodo.7019194.

48. A. Gretton, K. Borgwardt, M. J. Rasch, B. Scholkopf, A. J. Smola, A kernel method for the two-sample problem. arXiv:0805.2368 [cs.LG] (2008).

49. P. K. Rubenstein, B. Schoelkopf, I. Tolstikhin, On the latent space of Wasserstein auto-encoders. arXiv:1802.03761 [stat.ML] (2018).

50. J. Gorham, L. Mackey, "Measuring sample quality with kernels," in *Proceedings of the International Conference on Machine Learning* (PMLR, 2017), pp. 1292–1301.

51. L. P. Kaelbling, M. L. Littman, A. W. Moore, Reinforcement learning: A survey. *J. Artif. Intell. Res.* **4**, 237–285 (1996). doi:10.1613/jair.301

52. R. S. Sutton, Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9–44 (1988). doi:10.1007/BF00115009

53. A. V. Ruban, S. Simak, P. A. Korzhavyi, H. L. Skriver, Screened Coulomb interactions in metallic alloys. II. Screening beyond the single-site and atomic-sphere approximations. *Phys. Rev. B Condens. Matter* **66**, 024202 (2002). doi:10.1103/PhysRevB.66.024202

54. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996). doi:10.1103/PhysRevLett.77.3865 Medline

55. B. Gyorffy, Coherent-potential approximation for a nonoverlapping-muffin-tin-potential model of random substitutional alloys. *Phys. Rev., B, Solid State* **5**, 2382–2384 (1972). doi:10.1103/PhysRevB.5.2382

56. J. Staunton, B. Gyorffy, A. Pindor, G. Stocks, H. Winter, The "disordered local moment" picture of itinerant magnetism at finite temperatures. *J. Magn. Magn. Mater.* **45**, 15–22 (1984). doi:10.1016/0304-8853(84)90367-6

57. F. Birch, Finite elastic strain of cubic crystals. *Phys. Rev.* **71**, 809–824 (1947). doi:10.1103/PhysRev.71.809

58. A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques To Build Intelligent Systems* (O'Reilly Media, ed. 2, 2019).

59. V. L. Moruzzi, J. F. Janak, K. Schwarz, Calculated thermal properties of metals. *Phys. Rev. B Condens. Matter* **37**, 790–799 (1988). doi:10.1103/PhysRevB.37.790 Medline

60. X. Li, S. Schönecker, X. Li, J. Zhao, L. Vitos, The influence of temperature on the elastic properties of body-centered cubic reduced activation steels. *Mater. Des.* **197**, 109282 (2021). doi:10.1016/j.matdes.2020.109282

61. J.-O. Andersson, T. Helander, L. Höglund, P. Shi, B. Sundman, Thermo-Calc & DICTRA, computational tools for materials science. *Calphad* **26**, 273–312 (2002). doi:10.1016/S0364-5916(02)00037-8

62. H. Saito, *Physics and Applications of Invar Alloys* (Maruzen, 1978).

63. G. Pupke, über eine Anisotropie der thermischen Ausdehnung bei Eisen-Niekel-Legierungen. *Z. Phys. Chem.* **207**, 91–110 (1957). doi:10.1515/zpch-1957-20711