



OPEN

On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets

Aaron Gilad Kusne^{1,2*}, Tieren Gao^{1*}, Apurva Mehta³, Liqin Ke⁴, Manh Cuong Nguyen^{4,5}, Kai-Ming Ho^{4,5}, Vladimir Antropov⁴, Cai-Zhuang Wang^{4,5}, Matthew J. Kramer⁴, Christian Long¹ & Ichiro Takeuchi¹

¹Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, USA, ²National Institute of Standards and Technology, Gaithersburg, MD 20899, USA, ³Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA, ⁴Ames Laboratory, Iowa State University, Ames, IA 50011, USA, ⁵Department of Physics and Astronomy, Iowa State University, Ames, IA, 50011, USA.

Received
11 June 2014Accepted
18 August 2014Published
15 September 2014Correspondence and
requests for materials
should be addressed to
I.T. (takeuchi@umd.
edu)* These authors
contributed equally to
this work.

Advanced materials characterization techniques with ever-growing data acquisition speed and storage capabilities represent a challenge in modern materials science, and new procedures to quickly assess and analyze the data are needed. Machine learning approaches are effective in reducing the complexity of data and rapidly homing in on the underlying trend in multi-dimensional data. Here, we show that by employing an algorithm called the mean shift theory to a large amount of diffraction data in high-throughput experimentation, one can streamline the process of delineating the structural evolution across compositional variations mapped on combinatorial libraries with minimal computational cost. Data collected at a synchrotron beamline are analyzed on the fly, and by integrating experimental data with the inorganic crystal structure database (ICSD), we can substantially enhance the accuracy in classifying the structural phases across ternary phase spaces. We have used this approach to identify a novel magnetic phase with enhanced magnetic anisotropy which is a candidate for rare-earth free permanent magnet.

Recent explosion of activities in computational materials design is rapidly changing the landscape of materials science, and massive databases of predicted materials properties are now generated out of high-throughput thermodynamic and electronic-structure calculations^{1–5}. To truly harness the benefits of potential gold mines contained in the large number of such materials “blue prints”, however, it is still necessary to complete the materials exploration cycle by experimentally verifying the computed results. To establish such a cycle with steady feedback to predictions, it is often desirable to also carry out experiments in a high-throughput manner.

High-throughput and combinatorial experimental methods have indeed become mainstream techniques in many arenas of materials science in recent years owing in part to development of sophisticated materials synthesis and rapid characterization techniques. Depending on the physical property, measurement and data acquisition speed are now matching or in some instances outpacing the library synthesis speed. In many cases, one can now collect the data faster than they can be analyzed. This disparity in data collection and analysis time is fueling interest in expanding the high-throughput materials science to include accelerated data analysis through the use of new machine learning algorithms.

In the past, machine learning techniques have been implemented to computed materials properties such as electronic structures^{6–17}. Recent applications in materials science include texture analysis in micrographs^{18–24}. However, compared to the widespread use in other domains such as medical diagnostics^{25,26}, computational finances^{27,28}, and natural language processing²⁹, by and large, applications of machine learning in materials science are in their infancy. This is particularly true for exploratory learning from experimental data, which has been mostly limited to simple latent variable analysis techniques such as principle component analysis. More effective data mining techniques are especially in dire need when a large amount of data are being generated in a turn-key manner and individual data sometimes takes on non-trivial, high-dimensional forms such as spectra or images, and thus data from combinatorial experimentation represents an ideal platform to carry out machine learning based analysis.

One demonstrated application of machine learning in data from combinatorial libraries is to employ algorithms to quickly separate different parts of libraries into groups. When this is performed on a large number of x-ray diffraction patterns taken across a composition spread, the algorithms can be used to guide and delineate



structural phase boundaries leading to rapid construction of a composition-structure relationship, often an end goal of many materials science experiments^{30–32}. The learned composition-structure relationship from diffraction data can also be extrapolated to relationships between structure and functional properties such as magnetostriction or piezoelectricity^{33,34}, with structural phase boundaries indicating potential regions of significant change in functional properties. Here, we have devised an efficient and computationally-inexpensive new algorithm based on mean shift theory (MST) which overcomes shortcomings of previous methods we had used (hierarchical clustering and non-negative matrix factorization), including the inability to handle diffraction peak shifting.

Equipping the data analysis platform with a means of accessing and utilizing data entries from existing archival databases can provide significant benefits: analyzing experimental data in rapid cross-referencing mode with verified/archived data entries can substantially increase reliability and one's confidence in analysis performance. To this end, for studying a large number of x-ray diffraction data, we have developed a technique to mix and guide clustering of experimental diffraction patterns with simulated diffraction patterns using entries from the Inorganic Crystal Structure Database (ICSD), containing over 166,000 known compounds.

Further integrating the analysis platform with direct access to data as it is being collected by the characterization systems can provide the added benefits of live data analysis. For example, when instrument time or financial constraints limit the number of measurements to be performed, live data analysis can provide a means for guiding one to the optimal (number of) samples to characterize, maximizing knowledge of the overall set of samples with a minimum number of measurements. This is of great utility to experiments being carried out, for instance, at synchrotron beamlines.

We have developed a comprehensive algorithm which allows on-the-fly analysis of diffraction data from combinatorial libraries as they are collected and rapidly cross-referenced with pre-selected entries from the ICSD. Below, we describe how this strategy was used to study a series of Fe-Co-X ternary thin film composition spreads (where X is a transition metal element such as Mo, W, Ta, Zr, Hf, and V)^{35,36} to explore novel magnetic phases with enhanced magnetic anisotropy. Through the use of mean shift theory³⁷ as the machine-learning technique of choice on-the-fly separation and grouping of hundreds of diffraction patterns taken at a synchrotron beamline have led to the identification of a Mo substituted Fe-Co, whose tetragonally distorted structure is the origin of enhanced magnetic anisotropy.

Unsupervised learning as applied to data from combinatorial materials libraries. Previously, hierarchical cluster analysis (HCA) has been used to sort samples within combinatorial libraries into groups with similar characteristics using x-ray diffraction patterns, Raman spectra, and ferroelectric hysteresis loops^{30,34}. In addition to separating the composition regions into groups with similar structural properties, comparison of clustering results from different measurement techniques has been effective in revealing hidden correlations between physical properties. It is important to recognize that however robust and automated such a procedure might become, the clustering results always require human inspection for consistency and agreement with physical rules, e.g. Gibb's phase rule. The high speed technique of principal component analysis and other multidimensional scaling methods have been used to take the high dimensional powder patterns, typically hundreds to thousands of dimensions long, and visualize them in a 2D or 3D space, allowing experts to quickly determine structural trends. The latent variable analysis tool of non-negative matrix factorization (NMF) has also been found to be useful for swiftly identifying different phase mixture regions through a rapid basis pattern deconvolution process³¹.

However, one problem these techniques have in common is that they do not perform as well in the presence of peak shifting (the movement of diffraction peaks as a function of continuously changing composition due to lattice constant change), and often cluster adjacent region with shifted lattice constants as different phases. Because shifting of peaks as a function of composition is also common in many other spectral data such as Raman and X-ray photoelectron spectroscopy, a robust algorithm which can keep track of moving peaks is needed for analysis of a variety of characterization techniques.

One solution to the peak shifting issue had been proposed in the method of dynamic time warping (DTW). This method provides a measure for identifying similarities between diffraction patterns despite the presence of peak shifts³⁸. LeBras, et al. incorporated the DTW measure into their clustering algorithm³⁹. The algorithm breaks down each diffraction pattern into a list of peaks and uses constraint programming to discover the constituent phases through hard constraints on the form of the constituent phases. This algorithm is resilient to peak shifting and enforces both phase region connectivity and the Gibbs' phase rule. However, it currently comes at a high computational cost. Evaluation time can run in the tens of hours, without additional data processing performed by the user – currently too slow for a turn-key operation of continuous measurement/analysis cycles.

Out of a number of relatively new machine learning techniques we have looked at⁴⁰, we have elected to implement MST for the current work. We have found that MST is resilient to the presence of background noise in diffraction patterns and peak shifting and provides good cluster connectivity. MST is also fast and computationally inexpensive, taking less than ten seconds to analyze the material systems discussed here, and, as such, can be used for high-throughput analysis⁴⁰. MST is a non-parametric density estimation based clustering method. While not a Bayesian method, MST's statistical framework provides a probabilistic perspective on cluster membership in a high speed algorithm. The density of samples in the sample feature space are assumed to be dictated by probability density functions (PDF) associated with underlying (and unknown) sample classifications. For the case of mapping a composition spread, the classifications are assumed to be phase mixture regions in the composition space. MST identifies the PDF modes and sorts samples into clusters associated with each mode. This method is robust and applicable to different phase distribution scenarios, i.e. the phase regions to be distinguished can be either phase pure or phase mixed regions. (Figure 1) The implementation of MST follows the work found in Ref 41.

The feature vectors used for this work are a concatenation of the relative ternary composition with x-ray diffraction pattern intensity in order of corresponding Q values. While only the two feature types of composition and x-ray diffraction patterns have been used for this analysis, the MST method in general is capable of clustering based on a larger set of features including scalar features such as coercive field and high dimensional features such as Raman spectra. We combine experimental sample vectors with simulated sample vectors derived from the ICSD to provide improved clustering results.

To validate the overall approach developed here, MST was applied to a set of diffraction data from a Fe-Ga-Pd composition spread³⁰, which were previously also analyzed with other clustering techniques in separating composition regions into groups with similar crystal structures (Figure 2).

Compared to hierarchical clustering (Fig. 2a) and non-negative matrix factorization (Fig. 2b), MST (Fig. 2d) provides similar clustering results separating composition regions into different structural phases previously identified.

Inclusion of ICSD indeed leads to improved clustering results when using a suboptimal choice of bandwidths (Fig. 2c and 2d). We therefore incorporate ICSD data in analysis of all systems when-

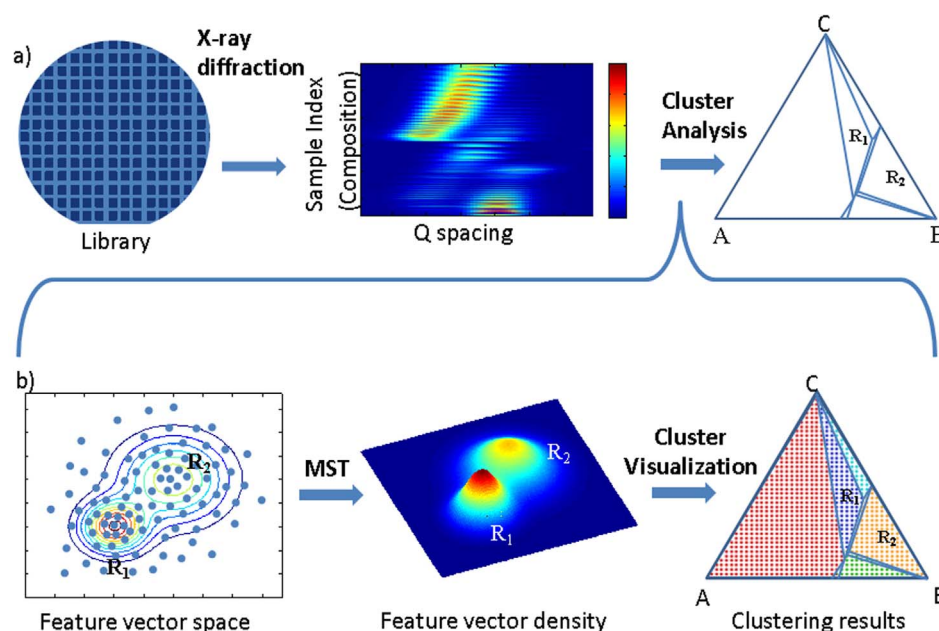


Figure 1 | Mean shift theory (MST) as applied to diffraction patterns taken from a composition spread wafer for rapid phase distribution analysis.

(a) X-ray diffraction data are taken from a thin-film composition spread wafer mapping a ternary (A-B-C) compositional phase diagram. The data are then analyzed using cluster analysis to produce a potential structural phase distribution diagram, identifying separated phase regions. (b) The MST process: 1) Feature vectors are produced for each sample on a combinatorial library. Each sample is projected into the feature vector space - shown here as 2 dimensional and unitless for ease of visualization, and the feature vector density is correlated to an underlying probability density function (PDF) for each 'hidden' classification, which in this case are assumed to be two separated different phase regions R1 and R2. 2) PDF analysis is performed using MST-based mode detection, and all samples from the same PDF are clustered together.

ever data are available. Because most ICSD entries are for binary compounds, as is the case with most ternary composition phase diagram studies, this works in such a way so that the mapping of ternary composition spreads is guided by known binary phases.

High-throughput search of Fe-Co-X with enhanced magnetic anisotropy. The utility of the above mentioned method was demonstrated in the high-throughput experimental search of novel rare-earth-free permanent magnets. We are systematically exploring Fe-Co-X composition spreads, where X is one of the 3d, 4d or 5d transition-metal elements. The outstanding challenge in the permanent magnet materials community is to design a magnetic crystal with sufficient magnetization and a large magnetic anisotropy without the use of inherently large spin-orbit coupling present in rare-earth elements⁴². One avenue to pursue such compounds is to insert elements such as W, Mo, or V in small atomic percentage to Fe-Co, so that hybridization of electronic states due to the finite spin-orbit coupling of 4d or 5d electrons with those of Fe and Co would lead to magnetic anisotropy. Another possibility is that small atomic percentage of inserted X element in the unit cell of cubic Fe-Co leads to tetragonal distortion resulting in magnetocrystalline anisotropy^{43–45}.

Fe-Co-X thin film composition spreads encompassing large ternary composition variation were made by co-sputtering and annealed in vacuum at 700°C for X = Mo, W, Ta, Zr, Hf, and V³⁵. Their magnetic properties were initially screened by a scanning MOKE (magneto-optical Kerr effect) set up, followed by more detailed hysteresis measurements using vibrating sample magnetometry (VSM) once composition regions which display indication of enhanced out-of-plane anisotropy were identified. Based on the shapes of the hysteresis loops, we can separate the composition regions into different groups as shown in Fig. 3(a), where hysteresis loops are overlaid on their approximate positions on the composition spread wafer. Such a quick visualization scheme is effective in rapidly assessing the land-

scape of how magnetism evolves as a continuous function of composition³⁴.

A composition region which consistently displays increased coercive field (which we take as the sign of anisotropy enhancement) in a single-phase-like out-of-plane hysteresis loop was observed in Fe-Co-Mo composition spreads (hysteresis loops plotted in red in Fig. 3(a)). This group of compositions was found to be centered around Fe₇₈Co₁₁Mo₁₁. Fig. 3(b) shows the typical out-of-plane and in-plane hysteresis loops of a Fe₇₈Co₁₁Mo₁₁ film. Other major characteristics seen in this visualization include 1) a broad composition region where two-phase-like out-of-plane hysteresis curves are observed indicating the presence of two co-existing magnetic phases (within each sample) with different coercive fields, and 2) a region with atom % concentration of Mo larger than 40%, where the films largely display paramagnetic behavior due to nanocrystalline and amorphous films.

The out-of-plane coercive field can also be mapped on the Fe-Co-Mo ternary spread as shown in Fig. 3(c). We see a clear transition of magnetization easy axis from out-of-plane to in-plane, and it is reflected in the out-of-plane coercive field map as a function of composition: the films of high Fe and low Mo concentration areas show higher perpendicular coercive field indicated by the red spots (corresponding to highest values in the coercive field measured here), and the films of high Co or high Mo concentration areas show minimal to no perpendicular coercive field from the blue spots where the magnetizations predominantly are lying in the film plane.

In order to streamline the process of determining the phase distribution and structural evolution across composition spreads, we carried out synchrotron x-ray diffraction at beam line 7-2 at Stanford Synchrotron Radiation Laboratory. An automated scanning stage is used to quickly step from spot to spot on the composition spread wafer, so that each diffraction image is acquired in 30 seconds per spot, and it takes roughly two hours to obtain mapping of the entire spread with approximately 200 spots using a Pilatus detector. The

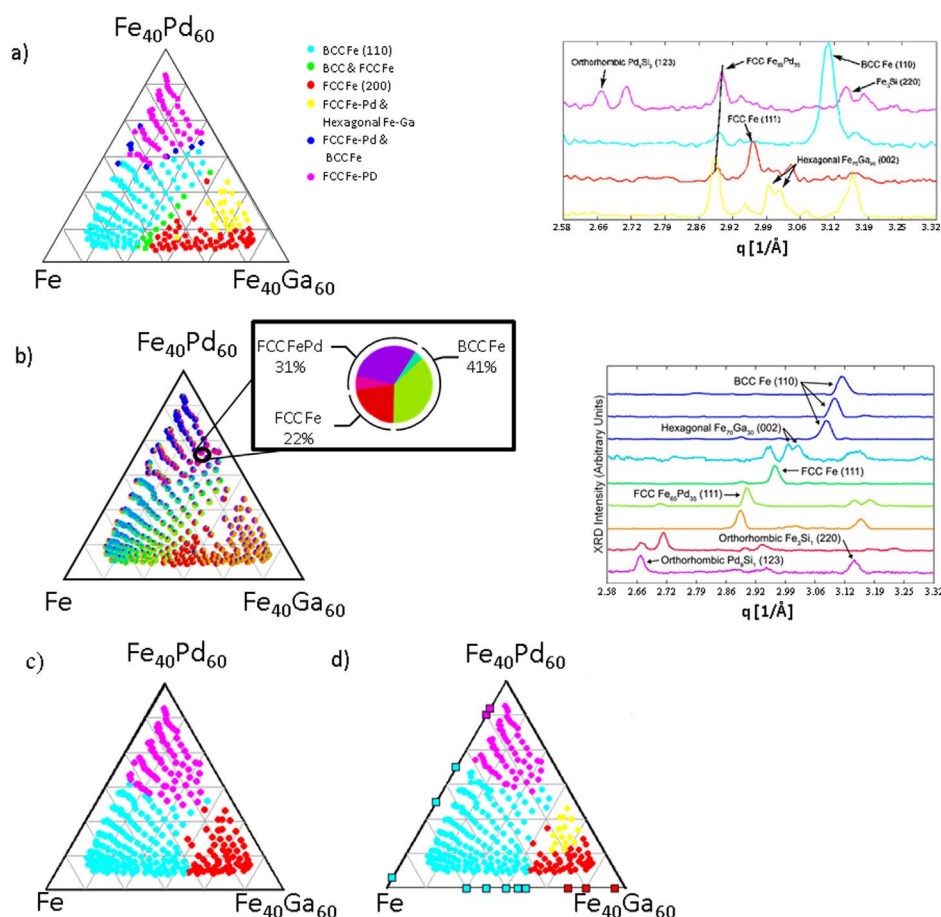


Figure 2 | Comparison of clustering results using different machine learning techniques applied to diffraction data taken from a Fe-Ga-Pd composition spread in the Fe-rich region. (a) Hierarchical cluster analysis clustering of diffraction pattern data (left); representative diffraction patterns from major cluster groups are plotted on the right; Reproduced from Ref. [30]. (b) Non-negative matrix factorization (NMF) analysis provides the positive principle components of the diffraction patterns. NMF gives deconvolution of patterns into components corresponding to peaks from pure phases, and accordingly, pie-charts are obtained at each average composition spot (left); peaks from basis patterns are then identified (right); Reproduced from Ref [31]. (c) High-speed mean shift theory based structural phase distribution analysis of experimental data using a suboptimal choice of bandwidth parameters. (d) High-speed MST results for experimental diffraction pattern data (ternary data points) and simulated data based on ICSD entries (square points along the binary lines) using the same suboptimal choice of bandwidths as in (c). ICSD based simulated data provides improved cluster stability despite suboptimal bandwidth choice.

data were analyzed on-the-fly using a combination of software, WxWindows Diffraction Integration Tool for integration of diffraction images into diffraction patterns, and CombiDM-Live for live clustering analysis. The fast and computationally-inexpensive MST implemented in this manner enabled the live analysis in such a way so that as the diffraction images are acquired for a spread wafer, updated clustering results are available for human inspection after measurement of each spot. Because of the robustness of the MST results, no noise smoothing was required for the analysis.

A quick inspection of the clustering behavior of the diffraction patterns using MST (Fig. 3 (d)) on a Fe-Co-Mo spread revealed that a group of compositions with enhanced single-phase-like coercive field (in Fig. 3(c)) coincides with a clustered region which displays a broad diffraction peak observed at $q = 3.06 \sim 3.09 \text{ \AA}^{-1}$ (red cluster in Fig. 3(d)). Fig. 3 (e) shows the intensity plot of all the X-ray diffraction patterns, grouped by MST in a selected q -region from 2.9 and 3.3 \AA^{-1} . From the top, they are: the region of interest with the broadened diffraction peak which coincides with the perpendicularly magnetized samples (red); the α -bcc Fe phase region with the sharp (110) peak (blue); the fcc Co phase, cubic Co-Mo or Fe-Mo phases (green). Because of the limited Mo solubility (at the annealing temperature of 700°C) and the non-equilibrium sputtering process,

the mapping of the observed phases is a non-equilibrium one. Some clustered regions (red and green) stretch into the high average Mo concentration region due partly to segregation of Mo (which we have confirmed with transmission electron microscopy (not shown)).

The films with the composition in the vicinity of $\text{Fe}_{78}\text{Co}_{11}\text{Mo}_{11}$ show perpendicular magnetic anisotropy with coercive field of 1.2 kOe (95.5 kA/m) and magnetization of 850 emu/cc (850 kA/m) (Fig. 3(b)). From these numbers, one can estimate the perpendicular anisotropy K_U using $H_K = \frac{2K_U}{M_S}$ assuming that the magnetization reversal is through coherent rotation. Here, M_S is the saturation magnetization, and H_K is the anisotropic field. The estimated perpendicular anisotropy K_U is typically around $3.6 \times 10^6 \text{ erg/cm}^3$ (27.0 \mu eV/atom) for $\text{Fe}_{78.4}\text{Co}_{10.8}\text{Mo}_{10.8}$ which is of the same order as that of the Co-Pt alloy⁴⁶.

To understand the origin of the enhanced coercive field and the out-of-plane anisotropy of this compound, we have employed the genetic algorithm (GA) search approach^{47,48} using the composition Fe_8CoMo , and the Scherrer equation was used in simulating the XRD patterns to take into account the broadening of the peak with a typical grain size of 50 nm (observed with transmission electron microscopy, not shown). There is good agreement between the

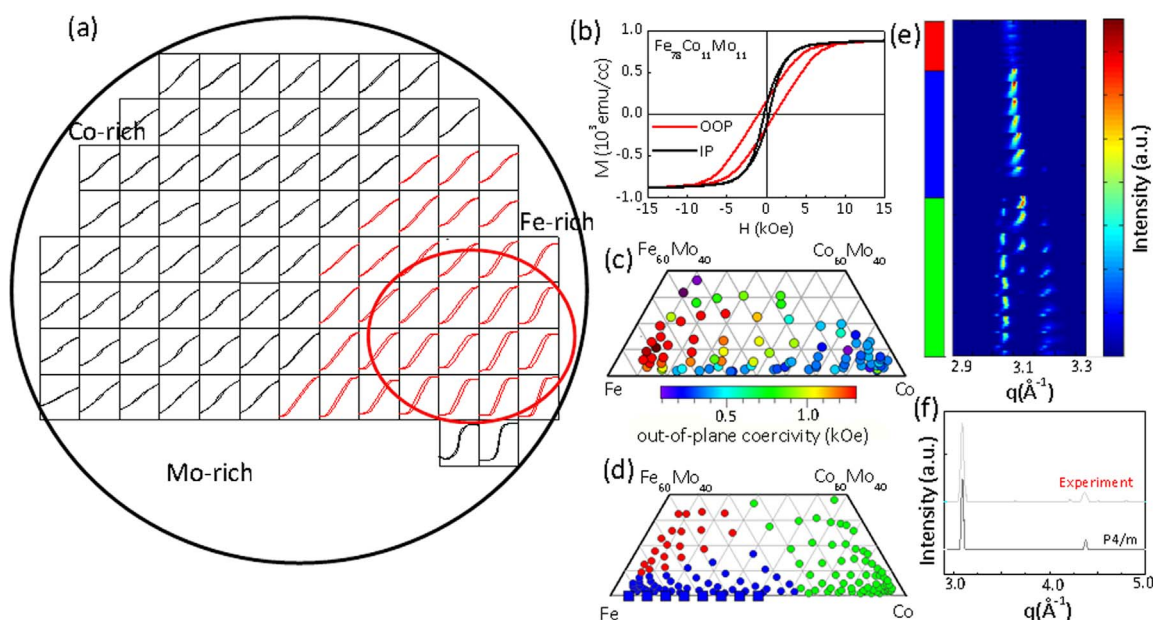


Figure 3 | Magnetic and structural property maps of Fe-Co-Mo composition spread: (a) out-of-plane hysteresis loops of Fe-Co-Mo samples for different compositions superimposed on the spread wafer positions; (b) typical out-of-plane (OOP, red) and in-plane (IP, black) hysteresis loops of $\text{Fe}_{78.4}\text{Co}_{10.8}\text{Mo}_{10.8}$ sample with perpendicular anisotropy ($1 \text{ emu/cc} = 10^3 \text{ A/m}$, $1 \text{ Oe} = 10^3/(4\pi) \text{ A/m}$); (c) Out-of-plane coercive field map of Fe-Co-Mo ternary alloys. (d) Clustering results of diffraction data using MST of Fe-Co-Mo ternary; (e) Intensity plot of x-ray diffraction patterns grouped by the clustering result (same color clusters as Fig. 3(d)); (f) Synchrotron X-ray diffraction spectrum (red) of $\text{Fe}_{78}\text{Co}_{11}\text{Mo}_{11}$ and calculated X-ray diffraction spectrum (black) with P4/m tetragonal structure.

experimental XRD pattern and the simulated pattern for a relatively low energy tetragonal P4/m structure (Fig. S1) identified by the GA search (Fig. 3(f)).

The formation energy of the P4/m structure is positive (7.7 meV/atom), indicating that the structure may be a metastable one. However, the structure can be stabilized in non-equilibrium processes depending on synthesis conditions. We believe it is not unreasonable that our co-sputtered films have resulted in formation of this structure. The magnetic moment of 4j-site Fe, 4k-site Fe, Co and Mo atoms in the P4/m Fe_8CoMo structure are 2.11, 2.55, 1.67 and $-0.61 \mu_B$, respectively, in our spin-polarized calculations. A non-collinear magnetic structure density function theory calculation including spin-orbit coupling interaction shows that P4/m Fe_8CoMo structure has a uniaxial easy axis along the (001) direction with magnetic anisotropy energy of $18.2 \mu\text{eV/atom}$, which compares well with the experimental observed value of $27.0 \mu\text{eV/atom}$.

Alternatively, the observed broadened bcc (110) peak can be taken as merging of two tetragonal peaks with a corresponding c/a ratio of approximately 1.04. We have also calculated the magnetocrystalline anisotropy energy (MAE) of Fe_8CoMo with a layered structure using this c/a ratio value using the full potential LMTO (linear muffin tin orbital) method. The calculation details can be found elsewhere⁴⁹. The calculated MAE is $28.0 \mu\text{eV/atom}$, which is very close to the experimentally extracted anisotropy value. Our calculations also indicate that a similar distortion may be brought about by substitution of other elements such as W and V.

We have thus experimentally realized a sought-after tetragonally distorted Fe-Co^{43,44} by introducing a third element and identified its composition. While the increase in anisotropy is modest, the fact that such a structure can be realized in a thin film form indicates that this is potentially a viable path toward manufacturing rare-earth free permanent magnets.

We have shown the use of the MST machine learning algorithm for on-the-fly analysis of x-ray diffraction and composition data toward the discovery of a novel rare-earth-free permanent magnet. While the MST method was found to be resilient to noise in the data,

specifically diffraction pattern noise, further algorithm developments should focus on Bayesian techniques that utilize data noise to quantify the uncertainty of cluster membership. Furthermore, an area of machine learning that may prove fruitful for high speed diffraction pattern analysis, including pure phase identification, is that of blind source separation: this is a research direction made popular by the famous cocktail party problem, where the goal is to identify individual speakers in a crowded room through the use of multiple simultaneous audio recordings, each recording containing a different mixture of the party goers' voices. For the case of the combinatorial library, the analogous goal is to identify pure phases from samples containing different phase mixtures. The diffraction patterns from the different phase mixture samples provide the multiple 'recordings' of the mixed pure phases to be analyzed with blind source separation. A collection of simple clustering methods such as MST and hierarchical cluster analysis can also be used in combination with a collection of blind source separation methods through the use of an ensemble algorithm that capitalizes on each method's benefits to provide an optimized analysis result. For example, the ensemble method could select the clustering result with the greatest cluster stability and then utilize the cluster boundaries to inform the blind source separation method on potential phase boundaries.

Naturally, the methods presented here can be extended to a variety of other spectral and image data. As previously demonstrated³⁴, such clustering algorithms (as well as visualization) can be applied simultaneously to data from a number of different physical properties measured on the same composition spread in order to rapidly delineate correlations between them. An important next step in developing the rapid materials exploration cycle is to directly integrate computational results and experimental results, as some computational databases provide extensive information on predicted stable crystal structures with lattice constants⁵⁰. Carrying out various machine learning algorithms on merged diffraction data from high-throughput experiments together with simulated diffraction patterns from predicted structures would provide instant feedback



on a number of computational results simultaneously. Such an integrated platform could thus truly speed up the process of materials discovery in the future.

Methods

For mean shift theory clustering, the concatenated sample vectors are first normalized by subtracting the vector mean. No noise smoothing is performed. The feature vectors in the composition space are assumed to be relatively uniformly distributed and the local density in both feature spaces is approximated using the normal kernel density estimator \hat{f} given by:

$$\hat{f} = \begin{cases} \sum_{i=1}^n \exp\left(-\frac{d_{L1}(x, x_i)^2}{2h^2}\right), & d_{L1}(x, x_i) \leq ch \\ 0, & d_{L1}(x, x_i) > ch \end{cases} \quad (1)$$

where d_{L1} is the L1 metric⁵¹, h is the bandwidth, and a symmetric kernel support of ch is used with $c = 1$. Two bandwidths are selected by the user, one for the diffraction data and one for the relative composition data, and these bandwidths determine the mode detection resolution. Choice of bandwidths impacts the number of clusters in the final clustering results, with suboptimal bandwidth choices resulting in either too many or too few clusters. The bandwidth selection heuristic used here and alternative methods for setting the bandwidths are discussed in the Supplementary Information.

The experimental sample vectors and the ICSD derived sample vectors are combined in the same feature space during clustering. The experimental samples are given a mass of 1 during MST mode detection, and all ICSD vectors are given the same user determined mass. Increasing the mass of the ICSD vectors increases local density in the feature space. In this way, the critically evaluated data from the ICSD can be given greater influence in guiding the clustering results than the experimentally obtained data. Cluster assignment is performed by associating each converged sample vector to the nearest original diffraction pattern, and all sample vectors that share the same associated diffraction pattern are assigned to the same cluster. Again, validation of clustering results against physical rules (e.g. Gibbs' phase rule for equilibrium systems) by a human expert is critical.

In incorporating data from the ICSD, the material structure data was exported in CIF format and then transformed into simulated diffraction patterns using Bruker's Topas software. Typical parameters were used and are listed in Table S1 in Supplementary Information. The simulated patterns from ICSD were then combined with their compositions to create simulated sample vectors. Most of the metallic film phases we are looking at deposited on Si/SiO₂ wafers predominantly grow with (110)/(101) texture (for α -Fe and related structures studied here), and therefore in generating the simulated patterns, (110) (and its family of) peaks were used as the main peaks.

The ternary composition spread films were deposited using magnetron co-sputtering on 3-inch diameter Si wafers (with 200 nm thick thermally oxidized SiO₂ layer) at room temperature. The details of the composition spread deposition technique are described elsewhere⁵². A physical shadow mask is used to separate the composition spread film into a grid of 4.5 mm \times 4.5 mm regions in order to characterize the structural and magnetic properties. The room-temperature deposited samples were then post-annealed in vacuum at 700 °C for 45 min. The average composition for both spreads of each thin film region is determined by wavelength dispersive spectroscopy (WDS). The synchrotron X-ray diffraction was performed with 16.5 keV energy at Beam line 7-2 at Stanford Synchrotron Radiation Laboratory (SSRL). Some of the Fe-Co-Mo wafers were cut into small pieces to characterize the magnetic and microstructural properties by vibrating sample magnetometry (VSM) and transmission electron microscopy (TEM).

The genetic algorithm (GA) searches were "from scratch" with only chemical compositions given. Structural optimizations during the GA searches were carried out by LAMMPS code⁵³ with Embedded Atom Method potential developed by Zhou et al.⁵⁴. The low energy structures obtained from GA searches were then fully relaxed again by first-principles calculations for the final structures. The first-principle spin-polarized density functional theory calculations⁵⁵ were performed by Vienna Ab initio Simulation Package⁵⁶ with PAW pseudopotential method^{57,58} within generalized-gradient approximation⁵⁹. The energy cutoff was 370 eV and the Monkhorst-Pack's scheme⁶⁰ was used for Brillouin zone sampling. A high-quality k-point grid, which is corresponding to 14 \times 14 \times 14 for body-centered cubic (bcc) Fe, was used in all calculations. The references for formation energy calculations were bcc Fe, hexagonal closed-pack (hcp) Co and bcc Mo. The X-ray diffraction patterns were simulated with X-ray wavelength $\lambda = 0.7517$ Å and the Scherrer equation was used for the peak broadening with sample size of 50 nm.

The magnetocrystalline anisotropy of layered Fe₃CoMo was calculated by using the full potential (FP) LMTO method with smoothed generalizations of Hankel functions for the envelope functions. Spin-orbit coupling has been added into the Hamiltonian and self-consistent calculations have been performed with spin quantization direction along different directions. The integration over the Brillouin zone was performed using the tetrahedron method with the Blöchl's correction. A dense k-mesh was used for integration to ensure a good accuracy of calculated MAE. Certain commercial equipment, instruments, or materials are identified in this publication for informational purposes only. Such identification is not intended to imply recom-

mendation or endorsement by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

1. Curtarolo, S. *et al.* The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
2. Jain, A. *et al.* A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* **50**, 2295–2310 (2011).
3. Klintonberg, M. The Electronic Structure Project. *Electron. Struct. Proj.* (2012). at <http://gurka.fysik.uu.se/esp-fs/> (Date of access: 01/05/2014).
4. Landis, D. D. *et al.* The Computational Materials Repository. *Comput. Sci. Eng.* **14**, 51–57 (2012).
5. Mihalkovič, M. & Widom, M. Ab initio calculations of cohesive energies of Fe-based glass-forming alloys. *Phys. Rev. B* **70**, 144107 (2004).
6. Fischer, C. C., Tibbetts, K. J., Morgan, D. & Ceder, G. Predicting crystal structure by merging data mining with quantum mechanics. *Nat. Mater.* **5**, 641–646 (2006).
7. Saad, Y. *et al.* Data mining for materials: Computational experiments with AB compounds. *Phys. Rev. B* **85**, 104104 (2012).
8. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S. & Ramprasad, R. Accelerating materials property predictions using machine learning. *Sci. Rep.* **3**, 1–6 (2013).
9. Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R. & Burke, K. Finding Density Functionals with Machine Learning. *Phys. Rev. Lett.* **108**, 253002 (2012).
10. Montavon, G. *et al.* Machine learning of molecular electronic properties in chemical compound space. *New J. Phys.* **15**, 095003 (2013).
11. Hautier, G., Fischer, C. C., Jain, A., Mueller, T. & Ceder, G. Finding Nature's Missing Ternary Oxide Compounds Using Machine Learning and Density Functional Theory. *Chem. Mater.* **22**, 3762–3767 (2010).
12. Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **108**, 058301 (2012).
13. Balabin, R. M. & Lomakina, E. I. Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. *J. Chem. Phys.* **131**, 074104 (2009).
14. Hansen, K. *et al.* Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies. *J. Chem. Theory Comput.* **9**, 3404–3419 (2013).
15. D' Avezac, M., Luo, J.-W., Chanier, T. & Zunger, A. Genetic-Algorithm Discovery of a Direct-Gap and Optically Allowed Superstructure from Indirect-Gap Si and Ge Semiconductors. *Phys. Rev. Lett.* **108**, 027401 (2012).
16. Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **13**, 17930–17955 (2011).
17. Meredig, B. *et al.* Combinatorial screening for new materials in unconstrained composition space with machine learning. *Phys. Rev. B* **89**, 094104 (2014).
18. Huffman, L., Simmons, J., De Graef, M. & Pollak, I. Shape priors for MAP segmentation of alloy micrographs using graph cuts. in *Proc. IEEE Stat. Signal Process. Workshop* 661–664 (2011).
19. Simmons, J. P. *et al.* Application and further development of advanced image processing algorithms for automated analysis of serial section image data. *Model. Simul. Mater. Sci. Eng.* **17**, 025002 (2009).
20. Comer, M., Bouman, C. A., Graef, M. D. & Simmons, J. P. Bayesian methods for image segmentation. *JOM* **63**, 55–57 (2011).
21. MacSleay, J. P., Simmons, J. P. & De Graef, M. On the use of 2-D moment invariants for the automated classification of particle shapes. *Acta Mater.* **56**, 427–437 (2008).
22. Gulsoy, E. B., Simmons, J. P. & De Graef, M. Application of joint histogram and mutual information to registration and data fusion problems in serial sectioning microstructure studies. *Scr. Mater.* **60**, 381–384 (2009).
23. Niezgoda, S. R., Yabansu, Y. C. & Kalidindi, S. R. Understanding and visualizing microstructure and microstructure variance as a stochastic process. *Acta Mater.* **59**, 6387–6400 (2011).
24. Kalidindi, S. R., Niezgoda, S. R. & Salem, A. A. Microstructure informatics using higher-order statistics and efficient data-mining protocols. *JOM* **63**, 34–41 (2011).
25. Lavrač, N. in *Artif. Intell. Med.* (Horn, W., Shahar, Y., Lindberg, G., Andreassen, S. & Wyatt, J.) 47–62 (Springer Berlin Heidelberg, 1999).
26. Cleophas, T. J. & Zwinderman, A. H. *Machine Learning in Medicine*. (Springer, 2013).
27. Boyarshinov, V. *Machine Learning In Computational Finance: Practical algorithms for building artificial intelligence applications*. (LAP LAMBERT Academic Publishing, 2012).
28. Kovalerchuk, B. & Vityaev, E. in *Data Min. Knowl. Discov. Handb.* (Maimon, O. & Rokach, L.) 1203–1224 (Springer US, 2005).
29. Jurafsky, D. & Martin, J. H. *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. (Pearson Prentice Hall, 2009).
30. Long, C. *et al.* Rapid structural mapping of ternary metallic alloy systems using the combinatorial approach and cluster analysis. *Rev. Sci. Instrum.* **78**, 072217 (2007).
31. Long, C., Bunker, D., Li, X., Karen, V. & Takeuchi, I. Rapid identification of structural phases in combinatorial thin-film libraries using x-ray diffraction and non-negative matrix factorization. *Rev. Sci. Instrum.* **80**, 103902 (2009).
32. Barr, G., Dong, W. & Gilmore, C. J. High-throughput powder diffraction. II. Applications of clustering methods and multivariate data analysis. *J. Appl. Crystallogr.* **37**, 243–252 (2004).



33. Hunter, D. *et al.* Giant magnetostriction in annealed Co_{1-x}Fe_x thin-films. *Nat. Commun.* **2**, 518 (2011).
34. Kan, D., Long, C. J., Steinmetz, C., Lofland, S. E. & Takeuchi, I. Combinatorial search of structural transitions: Systematic investigation of morphotropic phase boundaries in chemically substituted BiFeO₃. *J. Mater. Res.* **27**, 2691–2704 (2012).
35. Gao, T. *et al.* Combinatorial exploration of rare-earth-free permanent magnets: Magnetic and microstructural properties of Fe-Co-W thin films. *Appl. Phys. Lett.* **102**, 022419 (2013).
36. Sourmail, T. Near equiatomic FeCo alloys: Constitution, mechanical and magnetic properties. *Prog. Mater. Sci.* **50**, 816–880 (2005).
37. Fukunaga, K. & Hostetler, L. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Inf. Theory* **21**, 32–40 (1975).
38. Baumes, L. A., Moliner, M., Nicoloyannis, N. & Corma, A. A reliable methodology for high throughput identification of a mixture of crystallographic phases from powder X-ray diffraction data. *CrystEngComm* **10**, 1321–1324 (2008).
39. LeBras, R. *et al.* in *Princ. Pract. Constraint Program.* 2011 508–522 (Springer, 2011).
40. Mueller, T., Kusne, A. G. & Ramprasad, R. Machine learning in materials science: Recent progress and emerging applications. *Rev. Comput. Chem.* (Accepted for publication).
41. Comaniciu, D. & Meer, P. Mean shift: A robust approach toward feature space analysis. *Pattern Anal. Mach. Intell. IEEE Trans. On* **24**, 603–619 (2002).
42. Kramer, M. J., McCallum, R. W., Anderson, I. A. & Constantinides, S. Prospects for Non-Rare Earth Permanent Magnets for Traction Motors and Generators. *JOM* **64**, 752–763 (2012).
43. Burkert, T., Nordström, L., Eriksson, O. & Heinonen, O. Giant Magnetic Anisotropy in Tetragonal FeCo Alloys. *Phys. Rev. Lett.* **93**, 027203 (2004).
44. Andersson, G. *et al.* Perpendicular Magnetocrystalline Anisotropy in Tetragonally Distorted Fe-Co Alloys. *Phys. Rev. Lett.* **96**, 037205 (2006).
45. Yildiz, F., Przybylski, M., Ma, X.-D. & Kirschner, J. Strong perpendicular anisotropy in Fe_{1-x}Cox alloy films epitaxially grown on mismatching Pd(001), Ir(001), and Rh(001) substrates. *Phys. Rev. B* **80**, 064415 (2009).
46. Weller, D., Brändle, H., Gorman, G., Lin, C.-J. & Notarys, H. Magnetic and magneto-optical properties of cobalt-platinum alloys with perpendicular magnetic anisotropy. *Appl. Phys. Lett.* **61**, 2726–2728 (1992).
47. Deaven, D. & Ho, K. Molecular geometry optimization with a genetic algorithm. *Phys. Rev. Lett.* **75**, 288–291 (1995).
48. Ji, M., Wang, C.-Z. & Ho, K.-M. Comparing efficiencies of genetic and minima hopping algorithms for crystal structure prediction. *Phys. Chem. Chem. Phys.* **12**, 11617–11623 (2010).
49. Ke, L., Belashchenko, K. D., van Schilfgaarde, M., Kotani, T. & Antropov, V. P. Effects of alloying and strain on the magnetic properties of Fe₁₆N₂. *Phys. Rev. B* **88**, 024404 (2013).
50. Curtarolo, S. *et al.* AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* **58**, 227–235 (2012).
51. Krause, E. F. *Taxicab Geometry: An Adventure in Non-Euclidean Geometry.* (Dover Publications, 1987).
52. Takeuchi, I. *et al.* Identification of novel compositions of ferromagnetic shape-memory alloys using composition spreads. *Nat. Mater.* **2**, 180–184 (2003).
53. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
54. Zhou, X., Johnson, R. & Wadley, H. Misfit-energy-increasing dislocations in vapor-deposited CoFe/NiFe multilayers. *Phys. Rev. B* **69**, 144113 (2004).
55. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
56. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11185 (1996).
57. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
58. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
59. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
60. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).

Acknowledgements

This project was funded by the beyond rare-earth permanent magnet project (U.S. DOE EERE under the Vehicle Technologies Program), NIST grant #70NANB10H207, and partially funded by the ARPA-E REACT program. We acknowledge valuable discussions with V.L. Karen, U.R. Kettner, T. Vanderah, R. McCallum and I. Anderson. We also acknowledge the computing resources supports from Oak Ridge Leadership Computing Facility (OLCF) via INCITE project. The work at Stanford Synchrotron Radiation Lightsources, SLAC National Accelerator Laboratory is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

Author contributions

The on-the-fly machine learning procedure was conceived and developed by A.G.K. and I.T. T.G. made the combinatorial libraries and did the magnetic characterization. T.G., A.M., M.J.K. and I.T. carried out the synchrotron diffraction experiment. L.K. and V.A. performed the density functional theory calculations. The genetic algorithm calculations were carried out by M.C.N., C.Z.W. and K.M.H. C.L. performed the non-negative matrix factorization on the Fe-Ga-Pd data. A.G.K., T.G. and I.T. wrote the manuscript. All authors discussed and commented on the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kusne, A.G. *et al.* On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Sci. Rep.* **4**, 6367; DOI:10.1038/srep06367 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Supplementary Information

On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets

Aaron Gilad Kusne^{*1,2}, Tieren Gao^{*1}, Apurva Mehta³, Liqin Ke⁴, Manh Cuong Nguyen^{4,5}, Kai-Ming Ho^{4,5}, Vladimir Antropov⁴, Cai-Zhuang Wang^{4,5}, Matthew J. Kramer⁴, Christian Long¹, and Ichiro Takeuchi¹

- 1. Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, USA**
- 2. National Institute of Standards and Technology, Gaithersburg, MD 20899, USA**
- 3. Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA**
- 4. Ames Laboratory, Iowa State University, Ames, IA 50011, USA**
- 5. Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA**

**These authors contributed equally to this work.*

S1. Mean shift theory based clustering

The mean shift theory algorithm implementation used for this study is based on that of Comaniciu and Meer¹. Each sample is represented by a feature vector concatenating the relative ternary composition values with the powder pattern intensities in order of corresponding q values. The vectors are normalized by subtracting the vector mean and dividing by the variable standard deviation. No noise smoothing is performed. Clusters with 4 or fewer members are integrated into the nearest cluster in composition space. All calculations were performed in 64bit MATLAB on a 3.1 GHz quad core machine with 12 GB of RAM running Windows 7.

The clustering resolution is set by the user through the selection of bandwidth parameters, one for the relative ternary composition and one for the diffraction pattern intensities. The user can select multiple values for each bandwidth and results are given for each pair of bandwidth values. Bandwidth values that are too small will result in an excessive number of clusters, the limit being that each sample is assigned to its own cluster. Bandwidth values that are too large will result in all samples being clustered together. For the FeGaPd system in Figure 2c and d, a composition bandwidth of 0.2 and a diffraction pattern bandwidth of 10^6 were chosen; for the FeCoMo system in Figure 3d and e, the composition and diffraction pattern bandwidths were 0.08 and 58,000 respectively. For both figures, an ICSD mass of 1 was used.

As mentioned in the Methods section, cluster assignment is determined by associating each converged sample vector with the nearest diffraction pattern. Those sample vectors that share the same associated diffraction pattern are then clustered together. This method assumes that the modes of each cluster fall near one of the original diffraction patterns. Other methods exist for cluster assignment. For example, the distance between each converged sample vector (in both composition and diffraction space) could be tabulated and all vectors separated by a length less than the bandwidths would be considered “connected” nodes in a graphical representation. Graph connectivity could then be used to identify cluster assignment, with all vectors connected through a path in the graph assigned to the same cluster. Such a method is not prone to uncertainty in cluster assignment.

A simple heuristic allows the user to narrow in on a reasonable choice of bandwidths within approximately three attempts. Both bandwidths are initially set to values that vary by order of magnitude and the clustering results are inspected, providing a general sense of the appropriate range for each bandwidth. The bandwidth ranges are then reduced, ‘zooming in’ to the bandwidth parameter values with an appropriate number of clusters. Other heuristics for bandwidth selection have been discussed in the literature^{1,2}. A future implementation of CombiDM will include an option to automatically optimize bandwidths.

Fixed, user supplied bandwidths provide a flexibility in clustering that allow for expert input in clustering optimization. Alternatively, if reduced user interaction is desired, a variable bandwidth mean shift algorithm^{2,3} can be used, removing the need for user selected bandwidths. These methods rely on a local estimation of sample density (e.g. through the use of k-nearest neighbors) in the feature space to adaptively set the bandwidth value.

There are various ensemble methods which can be used to combine the results of mean shift theory with that of other clustering methods. One method would be to run each clustering method on the data and then select the optimal clustering result based on a cluster evaluation metric. For example, the cluster with the greatest cluster stability measure could be selected.

S2. ICSD based simulated diffraction patterns

When using data from the ICSD, the materials structure data was exported in CIF format and transformed into simulated diffraction patterns using Bruker’s Topas software. Typical x-ray parameters used and are listed in Table S2. The Lorentz-polarization correction factor was set to 0 for polarized radiation and 90 for unpolarized radiation. The texture parameter⁴ was set to 0.7 for the [1 1 0] direction. The ICSD based diffraction patterns intensities were evaluated at the same 2θ (or q) values as the experimental diffraction patterns.

Primary radius [mm]	Secondary radius [mm]	Detector slit width [mm]	Tube filament length in axial plane [mm]	Sample length in axial direction [mm]	Receiving slit length in axial plane [mm]	Number of rays from point X-ray source in axial plane
217.5	217.5	0.2	12	15	12	30

Table S1. Topas diffraction pattern simulation parameters

S3. Crystal structure of P4/m Fe₈CoMo

The P4/m Fe₈CoMo structure (Fig. S1) is a tetragonal structure containing 1 formula unit with lattice parameters Wyckoff positions shown in Table S2. In this structure, all nearest neighbors of the Co atom are 4*k*-site Fe atoms and those of the Mo atom are 4*j*-site Fe atoms. Even though the differences in the atomic radii of Fe, Co and Mo are small, they can lead to slight broadening of the (110) diffraction peak as observed in the experiment.

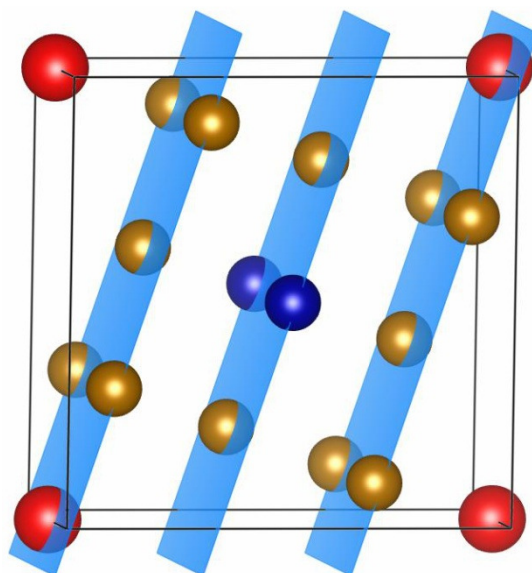


Figure S1. The crystal structure of P4/m Fe₈CoMo with (110) planes of the original bcc Fe indicated in light blue planes. The golden, blue and red balls are Fe, Co, and Mo atoms, respectively.

Parameters (Å)	Atom	x	y	z	M(μ_B)
	Fe (4j)	0.68971	0.10091	0.00000	2.11
a = 6.4295	Fe (4k)	0.20417	0.60188	0.50000	2.55
c = 2.8840	Co (1c)	0.50000	0.50000	0.00000	1.67
	Mo (1b)	0.00000	0.00000	0.50000	-0.61

Table S2. The lattice parameters, atomic position and the magnetic moment of each atom of P4/m Fe₈CoMo structure

S4. Layered structure of Fe₈CoMo

By using the full potential (FP) LMTO method, we have also calculated the magnetocrystalline anisotropy of Fe₈CoMo with 5 different unit-cell structures using the experimental c/a ratio of 1.04. Figure S2 shows the schematic of five different structures where the Fe, Co and Mo atoms are represented by the golden, blue and red balls, respectively. The calculation results are shown in Table S3. The results indicate that the layered structure #5 has the lowest energy and highest anisotropy of 28.0 $\mu\text{eV}/\text{atom}$ which is close to the value extracted from the experiment (27.0 $\mu\text{eV}/\text{atom}$). In this structure, the Co layers and Mo layers are sandwiched between Fe layers and the Mo layers tend to be separated from Co layers to lower the energy.

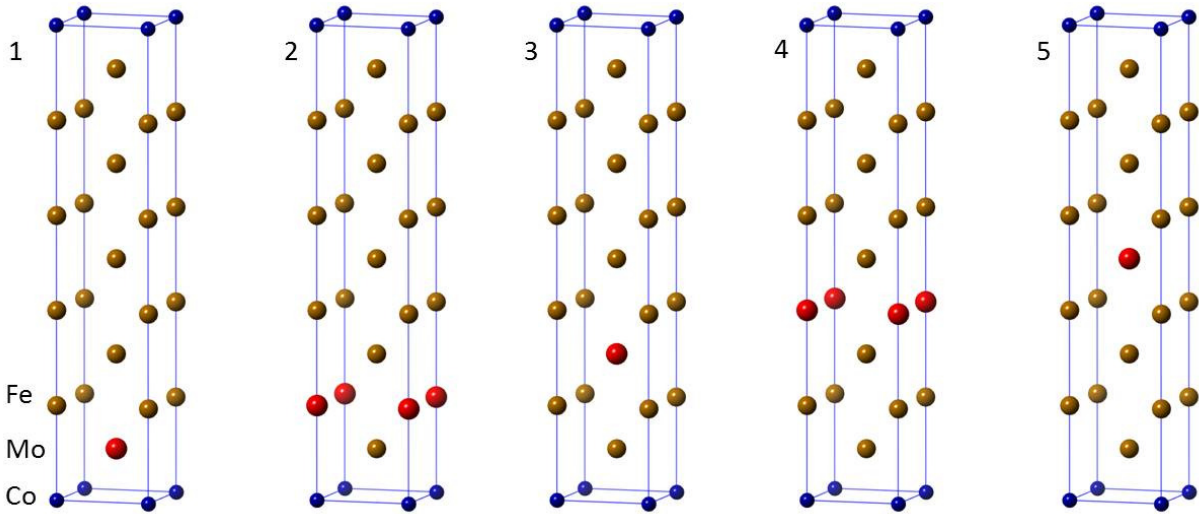


Figure S2. Five different unit cells of Fe₈CoMo with layered structures

Configuration	1	2	3	4	5
Energy(eV)	0.2474	0.1653	0.0312	0.0056	0
M(μ_B /cell)	19.79	20.12	20.22	20.41	20.38
K(μ eV/atom)	-49.3	12.1	13.9	18.5	28

Table S3. The formation energy, magnetic moment and anisotropy constant of five different layered structures Fe₈CoMo

REFERENCES

1. Comaniciu, D. & Meer, P. Mean shift: A robust approach toward feature space analysis. *Pattern Anal. Mach. Intell. IEEE Trans. On* **24**, 603–619 (2002).
2. Comaniciu, D. An algorithm for data-driven bandwidth selection. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 281–288 (2003).
3. Comaniciu, D., Ramesh, V. & Meer, P. The variable bandwidth mean shift and data-driven scale selection. in *Comput. Vis. 2001 ICCV 2001 Proc. Eighth IEEE Int. Conf. On* **1**, 438–445 (IEEE, 2001).
4. March, A. Mathematische Theorie der Regelung nach der Korngestalt bei affiner Deformation. *Z Krist.* **81**, 285–297 (1932).