

<https://doi.org/10.1038/s41524-025-01851-8>

# Networking autonomous material exploration systems through transfer learning

Check for updates

Naoki Yoshida<sup>1</sup>, Yutaro Iwabuchi<sup>1,2</sup>, Yasuhiko Igarashi<sup>1,3,4</sup> & Yuma Iwasaki<sup>4</sup>✉

Autonomous material exploration systems that integrate robotics, material simulations, and machine learning have advanced rapidly in recent years. Although their number continues to grow, these systems currently operate in isolation, limiting the overall efficiency of autonomous material discovery. In analogy to how human researchers advance materials science by sharing knowledge and collaborating, autonomous systems can also benefit from networking and knowledge exchange. Here, we propose a framework in which multiple autonomous material exploration systems form a network via transfer learning, selectively utilizing relevant knowledge from other systems in real time. We demonstrate this approach using three distinct autonomous systems and show that such networking significantly enhances the efficiency of material discovery. Our results suggest that the proposed framework can enable the development of large-scale autonomous material exploration networks, ultimately accelerating progress in material development.

The discovery of new materials is a major driver of technological innovation, enabling transformative advances across diverse industries. However, recent progress in materials science has led to increasingly complex material structures and compositions, resulting in an exponentially growing search space for novel materials. As a result, the process of identifying promising materials has become increasingly challenging. Traditionally, researchers have explored this vast search space manually, evaluating each candidate material one by one. Figure 1a shows the conventional human-driven material exploration loop, where materials are synthesized manually, their properties are characterized experimentally, and the resulting data are analyzed to guide the selection of the next material to be synthesized. Through repeated iterations of this loop, researchers can accumulate knowledge that can eventually lead to the discovery of new materials. However, this conventional approach has become inadequate for efficient material exploration as the search space continues to expand<sup>1</sup>.

To address these challenges, autonomous material exploration systems that integrate robotics, material simulations, and machine learning, particularly active learning, have garnered increasing attention in recent years<sup>2</sup>. These systems are broadly categorized into two types. The first type, shown in Fig. 1b, is an autonomous robotic material exploration system<sup>3</sup>. In this approach, robotic platforms automatically perform material synthesis and property characterization, whereas machine learning algorithms are used to

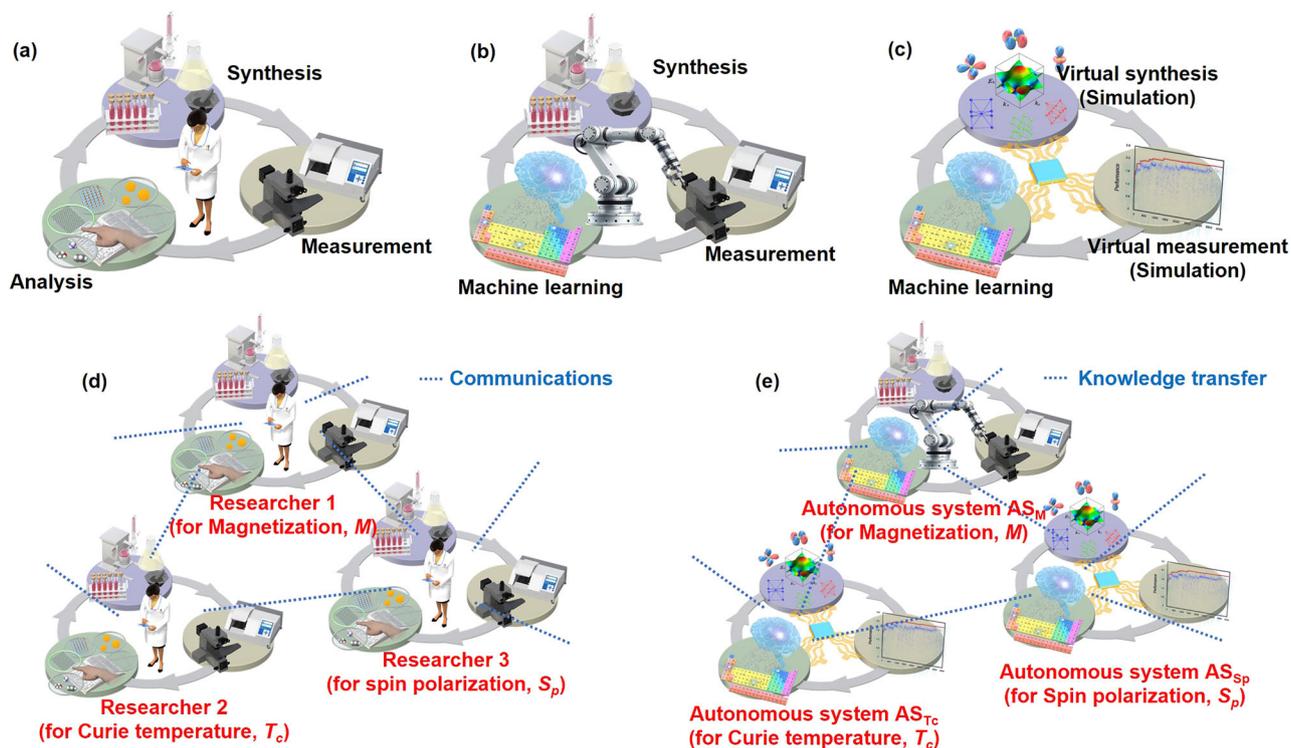
analyze the resulting data to determine the next candidate material. This process forms a closed-loop cycle driven by data and experimentation, enabling continuous exploration without human intervention. Notably, this approach has accelerated the discovery of high-performance materials. For example, it has been demonstrated that a mobile autonomous robot can autonomously perform synthesis, characterization, and data analysis, successfully discovering new photocatalytic materials<sup>4</sup>. Several other autonomous systems have been developed and are actively being used to facilitate the discovery of novel materials<sup>5–17</sup>.

The second approach, as shown in Fig. 1c, involves simulation-based autonomous material exploration systems. For example, density functional theory (DFT) simulations can be regarded as virtual material synthesis and characterization, where a material's composition and structure are used as inputs, and its physical properties are generated as outputs. By coupling such simulations with machine learning in a closed-loop framework, materials can be explored autonomously *in silico*, ultimately leading to the discovery of candidates with desirable properties. One notable example is a simulation-based autonomous system that combines DFT calculations with machine learning to explore the large chemical space of magnetic materials, successfully identifying a new alloy with exceptionally high magnetization ( $M$ )<sup>18</sup>. Many other simulation-based systems have also been developed and are actively advancing the discovery of materials computationally<sup>19–28</sup>.

<sup>1</sup>Degree Programs in Systems and Information Engineering, Graduate School of Science and Technology, University of Tsukuba, Tsukuba, Ibaraki, Japan.

<sup>2</sup>Research Center for Magnetic and Spintronic Materials (CMSM), National Institute for Materials Science (NIMS), Tsukuba, Japan. <sup>3</sup>Tsukuba Institute for Advanced Research, Tsukuba, Ibaraki, Japan. <sup>4</sup>Center for Basic Research on Materials (CBRM), National Institute for Materials Science (NIMS), Tsukuba, Japan.

✉ e-mail: IWASAKI.Yuma@nims.go.jp



**Fig. 1 | Autonomous material exploration systems and its networking.** **a** Schematic of the conventional human-driven material exploration loop, including material synthesis, property characterization, and decision-making for the next candidate material. **b** Robotic autonomous material exploration system with robotic

experimentation and machine learning. **c** In silico autonomous material exploration system with material simulation and machine learning. **d** Illustration of knowledge exchange among various researchers for material discovery. **e** Illustration of knowledge exchange among various autonomous material exploration systems.

As previously described, autonomous material exploration systems that integrate robotics, simulations, and machine learning are being actively developed worldwide and are operating in both physical and virtual environments. The applicability of these systems is expected to expand further across a vast material search space, contributing to the discovery of numerous novel materials.

Despite this progress, current robotic and simulation-based systems typically operate in isolation, which substantially limits their overall efficiency. To visualize this limitation, we consider the conventional human-driven approach to materials research. Figure 1d shows three researchers, each focusing on a distinct target property out of magnetization ( $M$ ), Curie temperature ( $T_c$ ), and spin polarization ( $S_p$ ).  $M$  is a fundamental property of magnetic materials that is crucial for applications such as permanent magnets and magnetic devices.  $T_c$  represents the thermal threshold at which a material transitions from a ferromagnetic to a paramagnetic state and is essential for ensuring thermal stability in magnetic systems.  $S_p$ , defined as the imbalance between up-spin and down-spin electrons at the Fermi level, is critical for spintronic applications, including magnetic sensors and data storage devices.

Each researcher independently performs iterative cycles of synthesis, property characterization, and data analysis to optimize the respective target properties. Notably, they also engage in discussions and exchange insights. For instance, a researcher investigating high- $M$  materials (Researcher 1) may benefit from insights shared by colleagues studying high- $T_c$  (Researcher 2) or high- $S_p$  materials (Researcher 3). By incorporating the trends observed in  $T_c$  or  $S_p$ , Researcher 1 can refine their experimental strategy, potentially improving discovery efficiency. This example highlights the importance of interdisciplinary knowledge exchange in human-driven research—an element currently lacking in most autonomous systems.

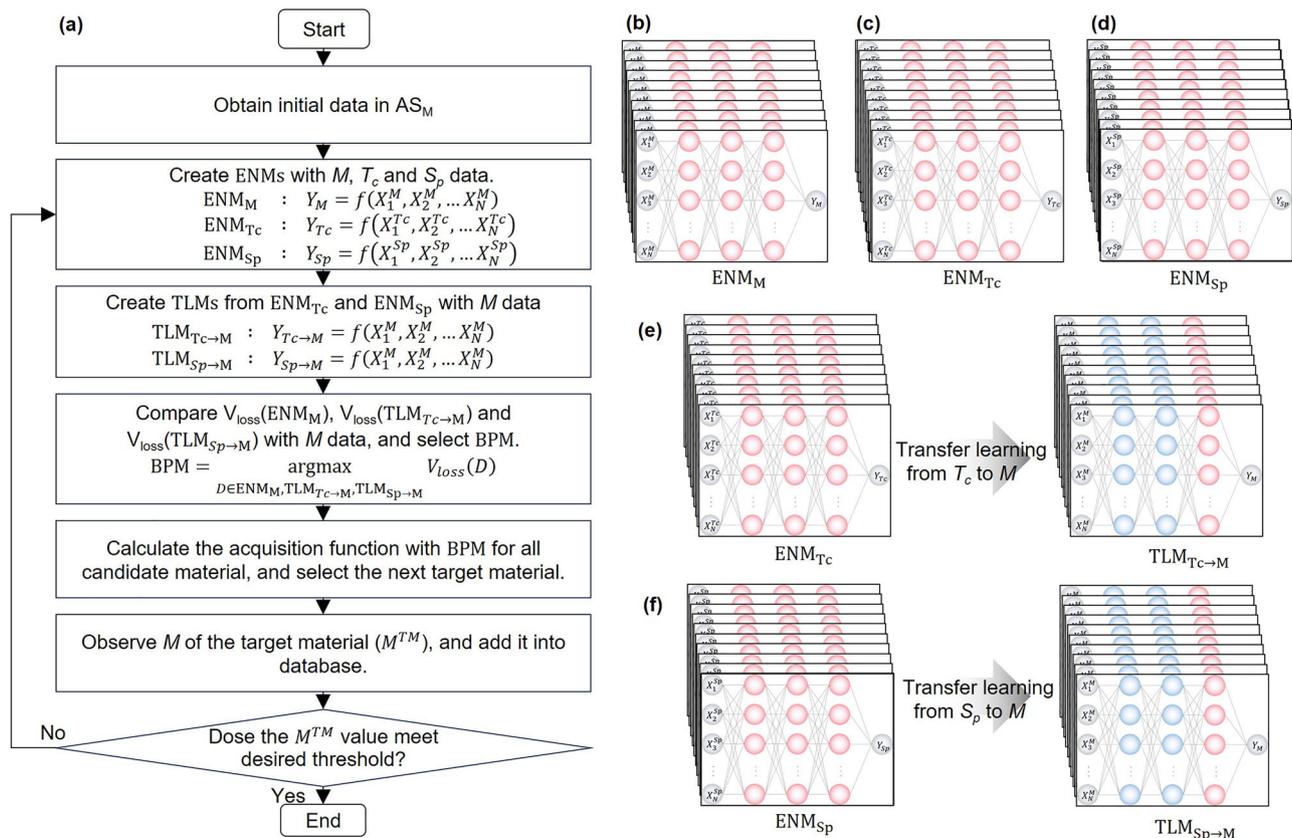
Unlike human researchers, current autonomous material exploration systems lack mechanisms for cross-domain knowledge sharing. Each system operates independently, relying solely on data related to its specific

target properties. For example, an autonomous system designed to discover materials with high  $M$ —referred to as the autonomous system for  $M$  ( $AS_M$ )—conducts Bayesian optimization using only  $M$ -related data. Even if this system receives additional data from other systems, such as  $T_c$  data from  $AS_{T_c}$  or  $S_p$  data from  $AS_{S_p}$ , integrating this external information into its predictive models remains challenging. A naïve approach may involve incorporating  $T_c$  or  $S_p$  as additional input features in  $AS_M$ 's prediction model for  $M$ . While this approach may improve predictive accuracy in certain cases, it limits the model's predictions to materials for which  $T_c$  or  $S_p$  values are already available. Consequently, the optimization process of  $AS_M$  becomes confined to the subset of materials previously explored by  $AS_{T_c}$  or  $AS_{S_p}$ , significantly reducing the system's ability to discover new materials.

Therefore, simply sharing raw data among autonomous systems is insufficient for improving exploration efficiency. Instead, a novel framework is required—one that enables multiple autonomous material exploration systems to form a collaborative network. Such a framework could enhance overall performance by utilizing external knowledge rather than directly exchanging raw data, as illustrated in Fig. 1e.

Transfer learning has potential to serve as a key technology of such collaborative networks by facilitating the transfer of knowledge across autonomous systems. It is a machine learning method in which knowledge obtained from one task is adapted for use in a different but related task, thereby improving learning efficiency<sup>29</sup>. In materials science, transfer learning has been widely adopted and is recognized as a powerful and promising approach<sup>30–33</sup>. However, its application within autonomous exploration systems has been scarcely investigated and remains in its infancy<sup>34</sup>.

This study proposes a method that enables knowledge sharing among multiple autonomous material exploration systems by combining ensemble neural networks (ENNs) with transfer learning. Instead of exchanging raw data, these systems transfer learned representations, enabling the incorporation of insights from other target domains. We demonstrate that this



**Fig. 2 | Workflow of the autonomous material search exploration system using ENN and transfer learning for networking and sharing knowledge.** **a** Schematic workflow of the autonomous material exploration system for  $M$  ( $AS_M$ ). **b–d** Architecture of ENNs for predicting  $M$ ,  $T_c$ , and  $S_p$ , respectively. **e**  $TLM_{T_c \rightarrow M}$

constructed by fine-tuning the base model  $ENM_{T_c}$  for predicting  $M$ . **f**  $TLM_{S_p \rightarrow M}$  constructed by fine-tuning the base model  $ENM_{S_p}$  for predicting  $M$ . The corresponding workflows for  $AS_{T_c}$  and  $AS_{S_p}$  are provided in Supplementary Informations S1 and S2.

approach significantly improves the efficiency of autonomous material exploration.

## Results

### Networking of autonomous material exploration systems

To evaluate the effectiveness of the proposed autonomous material exploration network, we demonstrate that the three systems,  $AS_M$ ,  $AS_{T_c}$ , and  $AS_{S_p}$ , depicted in Fig. 1e, can successfully share knowledge through transfer learning. This networked configuration improves the efficiency of material exploration compared with isolated operations.

Figure 2a outlines the algorithm used by  $AS_M$  within the autonomous material exploration network. Initially, the  $AS_M$  observes the  $M$  values of a small set of candidate materials (10 in this study) to construct an initial dataset. This design reflects the practical reality that autonomous material exploration often begins with very limited prior data; accordingly, we intentionally kept the initial dataset small. Based on this dataset, ensemble neural network models (ENMs) were constructed. Ensemble learning improves prediction accuracy and enables the estimation of predictive uncertainty (variance)<sup>35,36</sup>. A variety of machine learning methods can be used for Bayesian optimization, including Gaussian Process Regression (GPR)<sup>37</sup>, the Sequential Model-based Algorithm Configuration (SMAC)<sup>38</sup>, and the Tree-structured Parzen Estimator (TPE)<sup>39</sup>. In this work, however, we adopted ENMs, which not only support Bayesian optimization but also provide a straightforward framework for incorporating transfer learning. Each of the 10 individual neural networks (NNs) in the ensemble consisted of three hidden layers containing 100, 100, and 10 neurons. Additional details regarding the ENM architecture and training procedures are provided in the “Methods” section.

Figure 2b shows the  $ENM_M$ , which was trained to predict  $M$  as the target variable  $Y_M$  using material descriptors ( $X_1^M, X_2^M, \dots, X_N^M$ ) as explanatory variables.

$$ENM_M : Y_M = f(X_1^M, X_2^M, \dots, X_N^M) \quad (1)$$

To ensure generalizability, we employed simple composition-based vectors with  $N$  components as material descriptors throughout this study. Because these descriptors depend only on stoichiometry, they can be applied to both robotic and simulation-based autonomous material exploration systems (Fig. 1b, c). Similarly, Fig. 2c shows the  $ENM_{T_c}$ , which was trained to predict  $T_c$  using the data obtained from  $AS_{T_c}$  and its corresponding material descriptors ( $X_1^{T_c}, X_2^{T_c}, \dots, X_N^{T_c}$ ).

$$ENM_{T_c} : Y_{T_c} = f(X_1^{T_c}, X_2^{T_c}, \dots, X_N^{T_c}) \quad (2)$$

Figure 2d shows the  $ENM_{S_p}$ , which was trained to predict  $S_p$  based on data acquired from  $AS_{S_p}$  and its associated material descriptors ( $X_1^{S_p}, X_2^{S_p}, \dots, X_N^{S_p}$ ).

$$ENM_{S_p} : Y_{S_p} = f(X_1^{S_p}, X_2^{S_p}, \dots, X_N^{S_p}) \quad (3)$$

Considering that the network consisted of three autonomous material exploration systems, three corresponding ENMs were constructed. Notably, the number of ENMs increases accordingly when additional autonomous systems are incorporated.

Subsequently, two transfer learning models (TLMs) were constructed to incorporate external knowledge into the prediction of  $M$  within  $AS_M$ .

Figure 2e shows  $TLM_{T_c \rightarrow M}$ , which transfers the knowledge and trends obtained in  $AS_{T_c}$  regarding  $T_c$  to  $AS_M$  for predicting  $M$ . Specifically, a pretrained model  $ENM_{T_c}$  trained on  $T_c$  was used as the base model. The first two hidden layers on the input side (blue) were frozen, whereas the output layer (red) was fine-tuned using the  $M$  data obtained from  $AS_M$ . This approach allows the model to utilize knowledge obtained from  $T_c$  to enhance  $M$  prediction, yielding the following TLM :

$$TLM_{T_c \rightarrow M} : Y_{T_c \rightarrow M} = f(X_1^M, X_2^M, \dots, X_N^M) \quad (4)$$

Similarly, Fig. 2f depicts  $TLM_{S_p \rightarrow M}$ , where  $ENM_{S_p}$  serves as the base model. The model was fine-tuned using the  $M$  data observed in  $AS_M$ .

$$TLM_{S_p \rightarrow M} : Y_{S_p \rightarrow M} = f(X_1^M, X_2^M, \dots, X_N^M) \quad (5)$$

Consequently, three models were obtained for predicting  $M$ : the non-transfer model  $ENM_M$  and the transfer learning models  $TLM_{T_c \rightarrow M}$  and  $TLM_{S_p \rightarrow M}$ .

Subsequently, validation loss  $V_{loss}$  was compared among  $ENM_M$ ,  $TLM_{T_c \rightarrow M}$ , and  $TLM_{S_p \rightarrow M}$  using the  $M$  data and material descriptors available in  $AS_M$ . In this evaluation, 80% of the available material data was used as training data, whereas the remaining 20% was used as validation data. The best-performing model (BPM) was then selected based on the lowest validation loss.

$$BPM = \underset{D \in \{ENM_M, TLM_{T_c \rightarrow M}, TLM_{S_p \rightarrow M}\}}{\operatorname{argmin}} V_{loss}(D) \quad (6)$$

This model-selection step is critical for avoiding negative transfer, an effect in which transfer learning may reduce prediction performance. If transfer learning proves beneficial, the BPM is either  $TLM_{T_c \rightarrow M}$  or  $TLM_{S_p \rightarrow M}$ ; otherwise, the non-transfer model  $ENM_M$  is adopted. As incorporating external knowledge does not always lead to improvement, this validation-based selection process ensures that only beneficial models are used.

The acquisition function is then calculated using the selected BPM to identify the next material candidate for  $M$  measurement. In this study, the upper confidence bound (UCB) criterion was employed as the acquisition function<sup>40</sup>. The  $M$  of the selected target material ( $M^{TM}$ ) was measured and

added to the dataset. If  $M^{TM}$  did not satisfy the desired threshold, then the models  $ENM_M$ ,  $ENM_{T_c}$  and  $ENM_{S_p}$  were retrained using the updated dataset, as shown in Fig. 2a. In this dataset, the amount of  $M$  data increased, along with that of  $T_c$  and  $S_p$  data, owing to the concurrent execution of  $AS_{T_c}$  and  $AS_{S_p}$  over time.

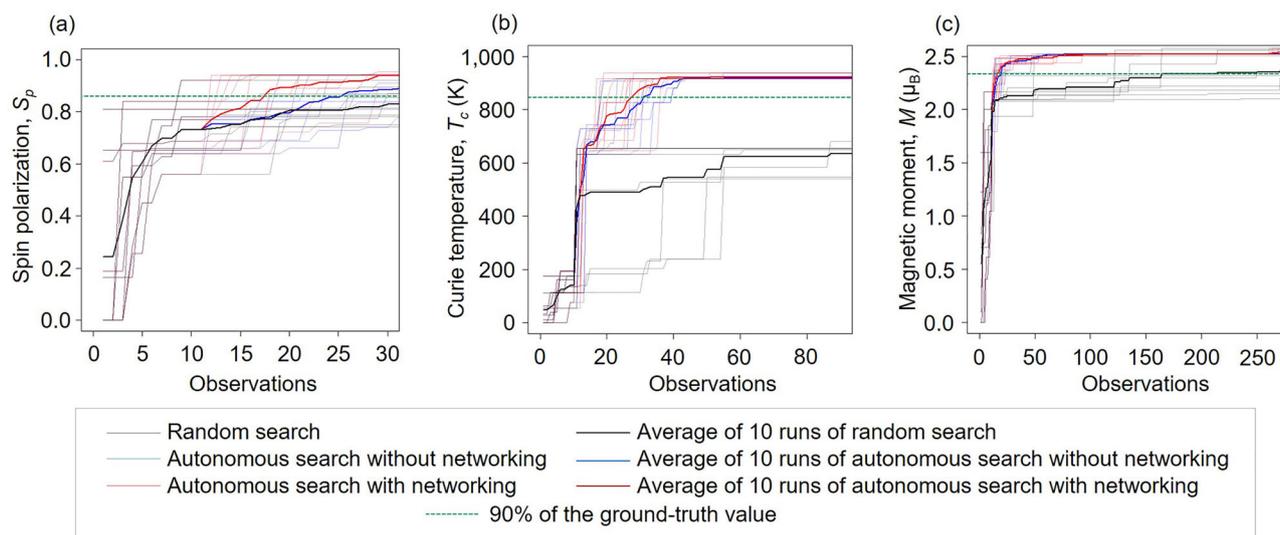
This iterative process enables  $AS_M$  to incorporate knowledge acquired from  $AS_{T_c}$  and  $AS_{S_p}$ , although these systems focus on different target properties ( $T_c$  and  $S_p$ ). Similarly,  $AS_{T_c}$  and  $AS_{S_p}$  can conduct autonomous exploration by utilizing knowledge from other systems. The details of the workflow and methodology from the perspectives of  $AS_{T_c}$  and  $AS_{S_p}$  are provided in Supplementary Information S1.

## Demonstration

To evaluate the effectiveness of the proposed autonomous material exploration network, a demonstration was conducted using a comprehensive dataset derived from high-throughput DFT calculations<sup>41</sup>. The dataset includes values for  $S_p$ ,  $T_c$ , and  $M$  for 16,908 ternary alloys with B2 crystal structures. These materials are composed of elements selected from a set of  $N = 38$ , including Li, Be, B, Mg, Al, Si, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Y, Zr, Nb, Mo, Ru, Rh, Pd, Ag, Cd, In, Sn, Sb, Hf, Ta, W, Ir, Pt, Au, and Pb.

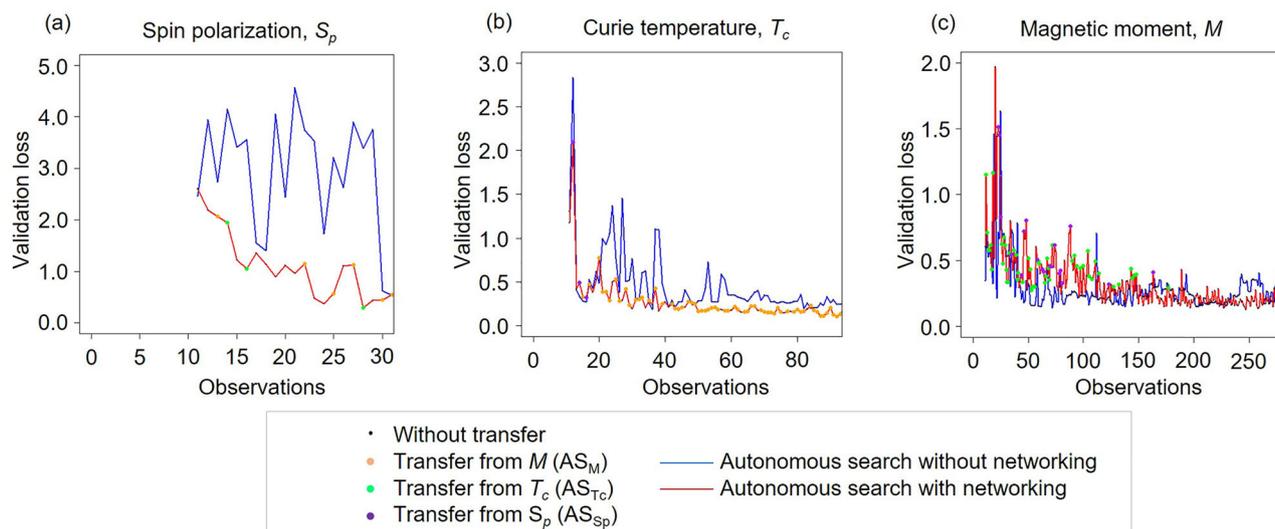
The number of initial observations was set to 10. The autonomous exploration speeds of  $AS_{T_c}$  and  $AS_{S_p}$  were adjusted relative to  $AS_M$  to reflect differences in the experimental complexity of measuring each target property. Specifically,  $AS_{T_c}$  was assigned a speed three times faster than  $AS_{S_p}$ , whereas  $AS_M$  was operated three times faster than  $AS_{T_c}$ . These relative speeds were designed to approximate the real-world feasibility of the model.  $M$  can be measured relatively quickly using techniques such as vibrating sample magnetometry and superconducting quantum interference devices<sup>42,43</sup>. By contrast, measuring the  $T_c$  requires an evaluation of the temperature dependence of  $M$ , which is time-consuming<sup>44</sup>. Measuring  $S_p$  is particularly challenging and typically involves complex experimental methods such as Andreev reflection<sup>45</sup>, nonlocal spin valve measurements<sup>46</sup>, and spin- and angle-resolved photoemission spectroscopy<sup>47</sup>.

Figure 3a shows the progression of the maximum observed  $S_p$  during autonomous exploration by  $AS_{S_p}$ . Light red lines represent results from 10 trials conducted with networking enabled via transfer learning from  $AS_{T_c}$  and  $AS_M$ , whereas the solid red line indicates the average performance. For comparison, results from the autonomous exploration using only local  $S_p$



**Fig. 3 | Comparison of the autonomous material exploration with and without networking using transfer learning.** a–c Progress of autonomous material exploration in  $AS_{S_p}$ ,  $AS_{T_c}$ , and  $AS_M$ , respectively. The light red and blue lines indicate the results of autonomous exploration with and without networking using transfer learning. The solid lines represent the average of 10 demonstration trials.

For comparison, results from a random search are shown in black. Green lines indicate 90% of the ground-truth values. While networking with transfer learning improves autonomous exploration efficiency in  $AS_{S_p}$  and  $AS_{T_c}$ , it has no noticeable effect in  $AS_M$ .



**Fig. 4 | Validation loss transition of BPM, and timing, frequency, and direction of transfer learning during autonomous exploration.** a–c Validation loss transition of the selected BPM in  $AS_{Sp}$ ,  $AS_{T_c}$ , and  $AS_M$ , respectively. The red and blue lines show the validation loss with and without networking using transfer learning, respectively. The selected BPMs are indicated as follows: green plots for the transfer

learning from  $AS_{T_c}$ , purple plots for the transfer learning from  $AS_{Sp}$ , orange plots for the transfer learning from  $AS_M$ , and small black plots for the non-transfer model. In  $AS_{Sp}$  and  $AS_{T_c}$ , networking improves the performance of the BPM, whereas no significant difference is observed in  $AS_M$ . This trend is consistent with the results of autonomous exploration efficiency shown in Fig. 3a–c.

data (i.e., without networking) are shown in blue, while those from a random search are shown in black. In the non-networked demonstration (blue line in Fig. 3), we trained three ENMs with distinct random initializations of the network weights and selected the BPM from among them. In addition, at the start of each exploration step, the network weights were randomly reinitialized before training. The networked approach (red) consistently outperformed both the non-networked (blue) and random (black) strategies, identifying high- $S_p$  materials more rapidly. On average, reaching 90% of the ground-truth value ( $S_p = 0.862$ ) required 17.4 observations with networking and 24.9 without.

Figure 3b presents the results for  $AS_{T_c}$ . Similar to  $AS_{Sp}$ , the networked approach demonstrated superior exploration efficiency compared with both the non-networked method and random search. The benefits of networking were particularly pronounced during the initial phase of autonomous exploration, where an improvement in efficiency was observed. Because the materials discovery process typically operates in a low-data regime and experimental or simulation throughput is limited, such early-stage gains are especially valuable. The mean number of observations required to reach 90% of the ground-truth value ( $T_c = 843.9$ ) was 27.4 with networking, compared with 31.8 without networking.

Figure 3c shows the corresponding results for  $AS_M$ . While both networked and non-networked autonomous strategies significantly outperformed the random search, no significant difference was observed between them. This suggests that networking has a limited impact on exploration efficiency in  $AS_M$ , at least under the current conditions. Reaching the ground-truth value ( $M = 2.352$ ) required, on average, 19.5 observations with networking and 20.2 without networking.

Overall, these results indicate that transfer-learning-based networking improves the efficiency of autonomous exploration, particularly for systems such as  $AS_{Sp}$  and  $AS_{T_c}$ . Notably, the model selection step prevents negative transfers by excluding models that degrade predictive performance. Consequently, networking does not reduce exploration efficiency. However, in some cases, such as  $AS_M$ , networking may not yield substantial improvements. A detailed analysis of the phenomenon is presented in the following section.

#### Validation loss transition on training data

We analyzed the transition of the validation loss during autonomous exploration, along with the timing, frequency, and direction of transfer

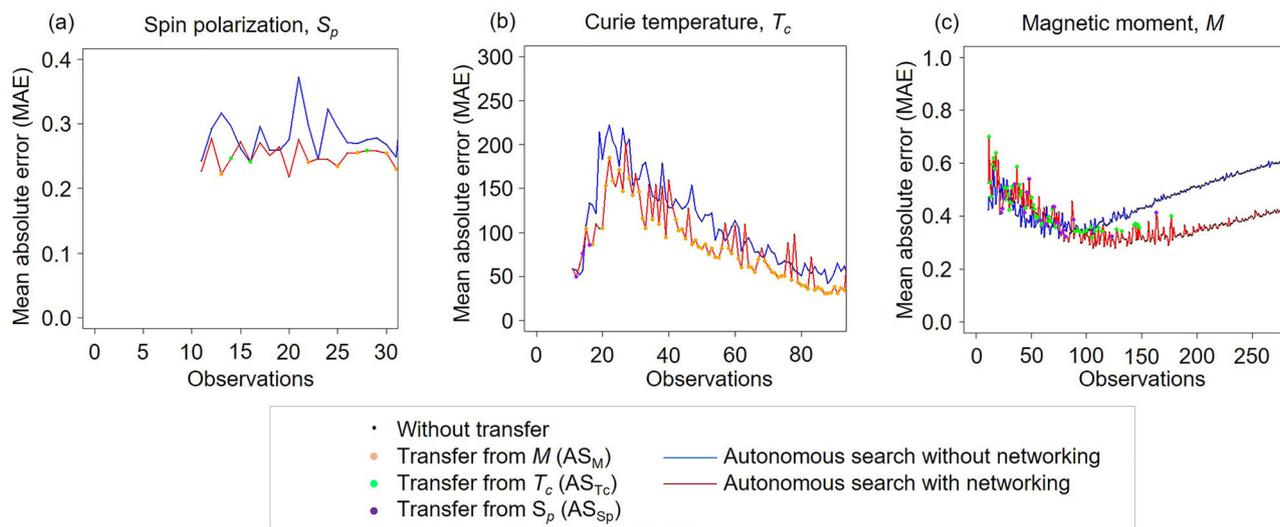
learning. Figure 4a shows the evolution of the validation loss associated with the BPM in  $AS_{Sp}$ , where the candidate models include  $ENM_{Sp}$ ,  $TLM_{T_c \rightarrow Sp}$ , and  $TLM_{M \rightarrow Sp}$ . The red curve represents the validation loss when networking via transfer learning is enabled, whereas the blue curve corresponds to a non-networked setting. Colored markers indicate the model selected as BPM at each step: orange for transfer from  $AS_M$  ( $TLM_{M \rightarrow Sp}$ ), green for transfer from  $AS_{T_c}$  ( $TLM_{T_c \rightarrow Sp}$ ), and small black dots for the non-transfer model ( $ENM_{Sp}$ ). The results show that  $AS_{Sp}$  effectively utilizes external knowledge from  $AS_M$  and  $AS_{T_c}$ , leading to a marked reduction in validation loss. This trend aligns with the enhanced exploration efficiency shown in Fig. 3a.

Figure 4b shows the validation loss trajectory of BPMs in  $AS_{T_c}$ , where the candidate models are  $ENM_{T_c}$ ,  $TLM_{M \rightarrow T_c}$ , and  $TLM_{Sp \rightarrow T_c}$ . Colored markers indicate the model selected as BPM at each step: orange for transfer from  $AS_M$  ( $TLM_{M \rightarrow T_c}$ ), purple for transfer from  $AS_{Sp}$  ( $TLM_{Sp \rightarrow T_c}$ ), and black for non-transfer model ( $ENM_{T_c}$ ). The transfer learning-enabled setting showed a lower validation loss than the non-networked setting. Notably, transfer from  $AS_M$  was frequently selected, suggesting that the knowledge gained by  $AS_M$ , owing to its higher exploration speed and larger data volume, was effectively transferred.

Figure 4c shows the validation loss trajectory of BPMs in  $AS_M$ , where the candidate models are  $ENM_M$ ,  $TLM_{T_c \rightarrow M}$ , and  $TLM_{Sp \rightarrow M}$ . The selected models are marked in green for transfer from  $AS_{T_c}$  ( $TLM_{T_c \rightarrow M}$ ), purple for transfer from  $AS_{Sp}$  ( $TLM_{Sp \rightarrow M}$ ), and black for the non-transfer model ( $ENM_M$ ). Unlike in the cases of  $AS_{Sp}$  and  $AS_{T_c}$ , no substantial improvements were observed in the transfer learning-enabled setting. This finding aligns with an earlier observation in Fig. 3c that networking has minimal impact on  $AS_M$ . In the later stages of exploration, the non-transfer model  $ENM_M$  was selected almost exclusively. This pattern likely reflects faster accumulation of  $M$  data by  $AS_M$  compared with the data collection rates of  $AS_{T_c}$  or  $AS_{Sp}$  for  $T_c$  and  $S_p$  data, resulting in superior performance of the local model  $ENM_M$  and limited benefit from external transfer, thus avoiding negative transfer.

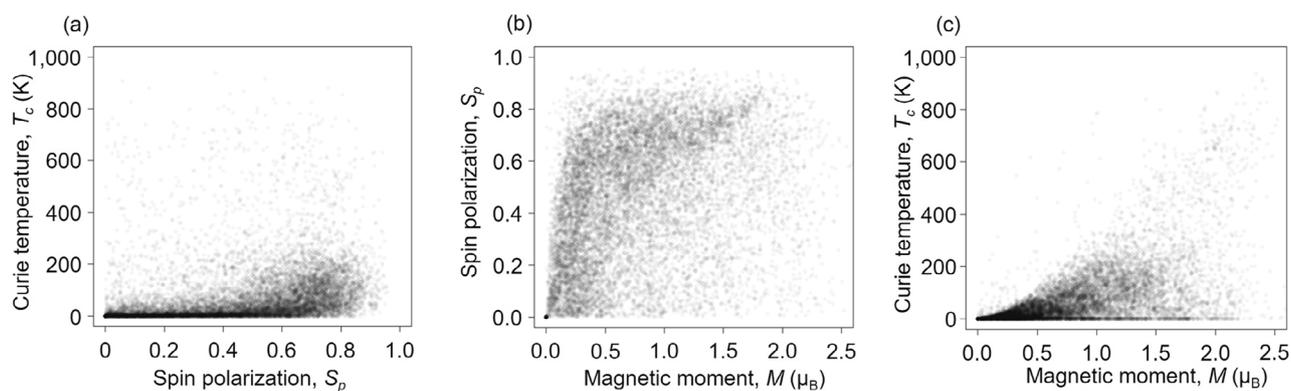
#### Generalization performance transition on unobserved data

We then evaluated the generalization performance of the BPMs for the unobserved materials within the search space, which consisted of 16,908 candidate materials. For instance, after 40 exploration steps conducted by  $AS_M$ , the dataset was divided as follows: 32 materials were used for training,



**Fig. 5 | Generalization performance transition of the BPM on unobserved data, and the timing, frequency, and direction of transfer learning during autonomous exploration.** **a–c** Generalization performance transition of the selected BPM in  $AS_{S_p}$ ,  $AS_{T_c}$ , and  $AS_M$ , respectively. The red and blue lines represent the validation loss with and without networking using transfer learning, respectively. The selected

BPMs are color-coded as follows: green plots for the transfer learning from  $AS_{T_c}$ , purple plots for the transfer learning from  $AS_{S_p}$ , orange plots for the transfer learning from  $AS_M$ , and small black plots for the non-transfer model. In all cases, networking improves the generalization performance of the BPM.



**Fig. 6 | Relationships among target properties  $M$ ,  $T_c$ , and  $S_p$ .** **a** Scatter plot of  $S_p$  versus  $T_c$ . **b** Scatter plot of  $M$  versus  $S_p$ . **c** Scatter plot of  $M$  versus  $T_c$ . Each plot reveals some degree of correlation, indicating that transfer learning is likely to be effective for these material properties.

8 for testing, and the remaining 16,868 were treated as unobserved. Notably, the generalization metrics computed using unobserved data were inaccessible during the actual autonomous exploration process and were used solely for retrospective evaluation in this study.

Figure 5a–c shows the evolution of generalization performance on unobserved materials during autonomous exploration by  $AS_{S_p}$ ,  $AS_{T_c}$ , and  $AS_M$ , respectively. As in Fig. 4a–c, the red and blue curves represent the generalization performance with and without networking via transfer learning, respectively. The colored markers denote the origin of the selected transfer models: orange for transfer from  $AS_M$ , purple from  $AS_{S_p}$ , and green from  $AS_{T_c}$ .

As shown in Fig. 5a, b, the networking approach yielded improved generalization compared with the non-networked setting. However, Fig. 5b reveals a temporary degradation in the generalization performance of  $AS_{T_c}$  as the number of exploration steps increases. This decline is a known artifact of Bayesian optimization and is commonly attributed to sampling bias<sup>48,49</sup>. For  $AS_{T_c}$ , the autonomous strategy favored materials with high  $T_c$  values, leading to a training set skewed toward a large  $T_c$ . The pool of unobserved materials included a substantial proportion of samples with low or zero  $T_c$ , resulting in a distribution mismatch that likely caused the observed drop in generalization performance.

As shown in Fig. 5c, the generalization performance of  $AS_M$  remained similar across the networked and non-networked settings during the early phase (fewer than 100 observations). Beyond this point, degradation due to training data bias became evident, similar to the case of  $AS_{T_c}$ . This degradation was mitigated more effectively in the networked scenario with transfer learning, suggesting its role in maintaining robustness against sampling bias in the later exploration stages.

## Discussion

Transfer learning is generally effective when the target properties exhibit underlying correlations. To examine this phenomenon, Fig. 6a–c present scatter plots of pairwise relationships among three material properties— $S_p$ ,  $T_c$ , and  $M$ —across the entire dataset. Figure 6a shows the relationship between  $S_p$  and  $T_c$ . Notably, materials with  $S_p$  values below approximately 0.6 tended to have  $T_c$  values near zero, whereas those with higher  $S_p$  values often exhibited finite  $T_c$ . Although a strong linear correlation was not observed, the distribution suggested a moderate dependence between  $S_p$  and  $T_c$ . Figure 6b shows the plot of  $M$  versus  $S_p$ . Despite considerable scatter, a general trend was observed. For  $M$  values below 0.5,  $S_p$  increased sharply with increasing  $M$ . Above this threshold, the rate of increase became more gradual, suggesting a weak but non-negligible relationship between  $M$  and

$S_p$ . Figure 6c shows the plot of  $M$  versus  $T_c$ , revealing a clear positive correlation. Additionally, the clustering of data points in the lower-right region indicates that  $M$  may be a limiting factor in achieving high  $T_c$  values. Overall, although no strong or universal correlations were observed, each pairwise comparison showed some degree of relationship, supporting the feasibility of applying transfer learning across these properties.

This study demonstrated that forming a network of autonomous material exploration systems via transfer learning enhanced the overall efficiency of material discovery. Notably, although networking enhances the efficiency of each autonomous exploration system, it never compromises or reduces their performance. When transfer learning introduced irrelevant external knowledge, the model selection mechanism defaulted to the non-transfer model. Consequently, knowledge sharing should be actively implemented in autonomous systems targeting seemingly unrelated materials or properties, and large-scale networks of such systems should be utilized to accelerate material discovery.

Moreover, analyzing the frequency, timing, and directionality of transfer learning events within the network enables a quantitative assessment of the strength of the intersystem relationships across different material properties. This analysis may yield unexpected insights. For instance, if a strong transfer relationship emerges between systems focused on distinct properties, it could motivate closer collaboration among researchers and communities working in these domains.

In this study, we introduced the concept of an autonomous material exploration network. Realizing such a system in practical deployments presents several challenges. We outline three of them below.

First, material descriptors are a key consideration. In this study, we employed simple composition-based vectors. Although such vectors are effective in many settings, they are not universal; although effective for inorganic systems, they often struggle with polymers. To improve exploration accuracy, it may be necessary to incorporate additional descriptors tailored to the exploration modality (e.g., robotic and simulation-based). Future work should focus on developing versatile descriptors that generalize across diverse environments and material classes, and on investigating transfer learning strategies across heterogeneous descriptor spaces.

Second, development aspects, such as communication protocols and metadata management, remain open. In our current knowledge-sharing framework, coordination among autonomous materials exploration systems can be achieved simply by exchanging trained models between systems. As experimental platforms and optimization algorithms evolve, however, it may become necessary to exchange and store richer metadata (e.g., data provenance, measurement conditions, hyperparameters, and model uncertainty). For future deployments, research should develop and evaluate reference implementations across a range of network topologies and establish communication protocols and interoperability standards.

Third, algorithmic choices warrant consideration. In this work, we demonstrated the concept using transfer learning with ENMs. Alternative machine-learning approaches, such as Gaussian process regression (GPR), Sequential Model-based Algorithm Configuration (SMAC), and Tree-structured Parzen Estimators (TPE), may further improve exploration efficiency. Knowledge sharing does not need to rely solely on transfer learning; for example, multitask neural networks may enable efficient sharing without explicit transfer learning. As autonomous materials exploration systems proliferate globally, frameworks and methods for knowledge sharing should be investigated systematically.

In summary, this study proposes a novel framework for accelerating material discovery by networking multiple autonomous material exploration systems that integrate robotics, material simulations, and machine learning. By utilizing ENNs and transfer learning, each system can extract knowledge from the data generated by other systems and apply it to improve its exploration efficiency. To validate this concept, we demonstrated a networked setup involving three autonomous exploration systems, each targeting distinct material properties: high  $M$ ,  $T_c$  and  $S_p$ . These systems were interconnected through a transfer learning-based network, enabling

dynamic knowledge sharing. The results confirmed that this networked approach improved exploration efficiency. Moreover, the proposed framework incorporated a model selection mechanism that automatically suppressed negative transfers from systems with insufficient data or weak interproperty correlations, ensuring that overall performance was not compromised. Future studies should consider analyzing the connections within large-scale networks of autonomous exploration systems, which may lead to the discovery of novel insights into the interfaces between different material systems. As autonomous material exploration systems continue to be used globally, the ability to connect and coordinate these systems will become increasingly important. The proposed approach provides a scalable and robust foundation for distributed, cooperative, and data-driven material exploration and offers a promising path for accelerating the discovery of novel functional materials.

## Methods

### ENNs

An ENN is a machine learning framework for quantifying predictive uncertainty by aggregating the outputs of multiple independently trained NNs. In this study, each constituent model in the ensemble is a feedforward NN (FNN) that receives material descriptors (i.e., elemental composition) as inputs and outputs a predicted material property, specifically,  $M$ ,  $T_c$  or  $S_p$ . The overall ENN prediction is obtained by averaging the outputs of the individual FNNs, whereas the predictive uncertainty is estimated as the variance across these outputs. By approximating a distribution over predictions, the ENN provides uncertainty estimates comparable in expressiveness to those of the Gaussian process regression<sup>35</sup>. Thus, ENNs can be effectively employed instead of Gaussian processes for tasks such as Bayesian optimization<sup>36</sup>.

In this study, each FNN comprised an input layer for material descriptors, three fully connected hidden layers with dimensions of 100, 100, and 10, and an output layer that returned the predicted property. All layers utilized a scaled exponential linear unit activation function. To improve the expressive capacity of the network, weight initialization was performed using a truncated normal distribution, as proposed by Klambauer et al. and LeCun et al.<sup>50,51</sup>.

Training was conducted using early stopping, with 20% of the data reserved as a validation set. The same validation strategy was applied during the selection process for the TLMs, as described in the subsequent section. The FNNs were implemented using the Keras application programming interface (API) in TensorFlow version 2.13.1, running in a Python 3.8.10 environment. The learning rate, batch size, and L2 regularization penalty were set as 0.0005, 8, and 0.001, respectively. The patience parameter for early stopping was set to 100 epochs.

### Transfer learning

Transfer learning is a technique that improves model performance by initially pretraining on a source dataset—often associated with a different but related task—and subsequently fine-tuning the model on a target dataset<sup>29–34</sup>. This approach enables the transfer of learned representations from non-target properties, which may differ from the optimization objective of the current autonomous exploration system.

In this study, transfer learning was applied by utilizing ENMs trained by other autonomous materials exploration systems. To preserve the learned features, the first two hidden layers (those closest to the input layer) of each source model were frozen, while the final hidden layer (closest to the output) was fine-tuned using the local target data. Using this approach, multiple ENMs were constructed, each corresponding to a different source system.

All transfer learning procedures were implemented using the Keras API within TensorFlow version 2.13.1 running on Python 3.8.10. The learning rate, batch size, and L2 regularization coefficient were set as 0.0005, 8, and 0.001, respectively. Early stopping was applied with a patience parameter of 100 epochs using 20% of the data as the validation set, consistent with the procedure described for ENN training.

## Acquisition function

The acquisition function is a central component of Bayesian optimization that guides the selection of the next candidate material to be evaluated. In this study, we adopted the UCB strategy<sup>40</sup>, which balances exploration and exploitation based on both predictive uncertainty and expected performance. Using the ENN, we computed the predicted mean and standard deviation of the target property across the entire search space. These values were then used to evaluate the UCB acquisition function, defined as:

$$UCB = \mu + \alpha \cdot \sigma \quad (7)$$

where  $\mu$  and  $\sigma$  denote the predicted mean and standard deviation for a given candidate material, and  $\alpha$  is a hyperparameter that governs the trade-off between exploration (favoring high uncertainty) and exploitation (favoring high predicted values). In this study,  $\alpha$  was fixed at 3.0, enabling the search algorithm to prioritize candidates predicted to either perform well or exhibit high uncertainty, thereby facilitating efficient exploration of the material space.

## Data availability

The dataset of B2-structured ternary alloys, including their magnetic moment, Curie temperature, and spin polarization values calculated by KKR-CPA (AkaiKKR), is available at the Materials Data Repository: <https://doi.org/10.48505/nims.5364>.

## Code availability

The code for networking of autonomous materials exploration systems is available at GitHub: <https://github.com/IgarashiLab/autonomous-material-network>.

Received: 30 June 2025; Accepted: 29 October 2025;

Published online: 09 December 2025

## References

- Ramprasad, R., Batra, R., Pilia, G., Mannodi-Kanakthodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
- Volk, A. A. & Abolhasani, M. Performance metrics to unleash the power of self-driving labs in chemistry and materials science. *Nat. Commun.* **15**, 1378 (2024).
- Szymanski, N. J. et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **624**, 86–91 (2023).
- Burger, B. et al. A mobile robotic chemist. *Nature* **583**, 237–241 (2020).
- Wang, C. et al. Autonomous platform for solution processing of electronic polymers. *Nat. Commun.* **16**, 1498 (2025).
- Pyzer-Knapp, E. O. et al. Accelerating materials discovery using artificial intelligence, high-performance computing and robotics. *npj Comput. Mater.* **8**, 84 (2022).
- Coley, C. W. et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, 6453 (2019).
- Nikolaev, P. et al. Autonomy in materials research: a case study in carbon nanotube growth. *npj Comput Mater.* **2**, 16031 (2016).
- Shimizu, R., Kobayashi, S., Watanabe, Y., Ando, Y. & Hitosugi, T. Autonomous materials synthesis by machine learning and robotics. *APL Mater.* **8**, 111110 (2020).
- Zhi, L. et al. Robot-accelerated perovskite investigation and discovery. *Chem. Mater.* **32**, 5650–5663 (2020).
- Roch, A. M. et al. ChemOS: orchestrating autonomous experimentation. *Sci. Robot* **3**, 19 (2018).
- Attia, P. M. et al. Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020).
- Granda, J. M., Donina, L., Dragone, V., Long, D.-L. & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377–381 (2018).
- Szymanski, N. J. et al. Toward autonomous design and synthesis of novel inorganic materials. *Mater. Horiz.* **8**, 2169–2198 (2021).
- Ren, Z., Ren, Z., Zhang, Z., Buonassisi, T. & Li, J. Autonomous experiments using active learning and AI. *Nat. Rev. Mater.* **8**, 563–564 (2023).
- Tamuda, R., Tsuda, K. & Matsuda, S. NIMS-OS: an automation software to implement a closed loop between artificial intelligence and robotic experiments in materials science. *Sci. Technol. Adv. Mater.* **3**, 2232297 (2023).
- Ament, S. et al. Autonomous materials synthesis via hierarchical active learning of interfacial energies. *Sci. Adv* **7**, eabg4930 (2021).
- Iwasaki, Y., Sawada, R., Saitoh, E. & Ishida, M. Machine learning autonomous identification of magnetic alloys beyond the Slater-Pauling limit. *Commun Mater* **2**, 1–7 (2021).
- Cubuk, E. D. et al. Accelerated discovery of stable materials with deep learning. *Nature* **624**, 80–85 (2023).
- Iwasaki, Y. et al. Autonomous search for half-metallic materials with B2 structure. *Sci. Technol. Adv. Mater.* **4**, 2403966 (2024).
- Jalem, R. et al. Bayesian-driven first-principles calculations for accelerating exploration of fast ion conductors for rechargeable battery application. *Sci. Rep.* **8**, 5845 (2018).
- Furuya, D., Miyashita, T., Miura, Y., Iwasaki, Y. & Kotsugi, M. Autonomous synthesis system integrating theoretical, informatics, and experimental approaches for large-magnetic-anisotropy materials. *Sci. Technol. Adv. Mater: Methods* **2**, 280–293 (2022).
- Iwasaki, Y., Jaekyun, H., Sakuraba, Y., Kotsugi, M. & Igarashi, Y. Efficient autonomous material search method combining ab initio calculations, autoencoder, and multi-objective Bayesian optimization. *Sci. Technol. Adv. Mater.* **2**, 365–371 (2022).
- Ma, X. Y. et al. High-efficient ab initio Bayesian active learning method and applications in prediction of two-dimensional functional materials. *Nanoscale* **13**, 14694–14704 (2021).
- Iwasaki, Y., Ogawa, D., Kotsugi, M. & Takahashi, Y. K. Autonomous materials search using machine learning and ab initio calculations for L10-FePt-based quaternary alloys. *Sci. Technol. Adv. Mater.* **5**, 2470114 (2025).
- Gubaev, K., Podryabinkin, E. V., Hart, G. L. W. & Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comput. Mater. Sci.* **156**, 148–156 (2019).
- Farache, D. E., Verduzco, J. C., McClure, Z. D., Desai, S. & Strachan, A. Active learning and molecular dynamics simulations to find high melting temperature alloys. *Comput. Mater. Sci.* **209**, 111386 (2021).
- Iwasaki, Y. Autonomous search for materials with high Curie temperature using ab initio calculations and machine learning. *Sci. Technol. Adv. Mater.* **4**, 2399494 (2024).
- Gupta, V. et al. Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data. *Nat. Commun.* **12**, 6595 (2021).
- Chang, R., Wang, Y. & Ertekin, E. Towards overcoming data scarcity in materials science: unifying models and datasets with a mixture of experts framework. *npj Comput. Mater.* **8**, 242 (2022).
- Minami, S. et al. Scaling law of Sim2Real transfer learning in expanding computational materials databases for real-world predictions. *npj Comput. Mater.* **11**, 146 (2025).
- Li, X. et al. A transfer learning approach for microstructure reconstruction and structure-property predictions. *Sci. Rep.* **8**, 13461 (2018).
- Yamada, H. et al. Predicting materials properties with little data using shotgun transfer learning. *ACS Cent. Sci.* **5**, 1717–1730 (2019).
- Hwang, J. & Iwasaki, Y. Improving efficiency of autonomous material search via transfer learning from nontarget properties. *Sci. Technol. Adv. Mater.* **3**, 2254202 (2023).
- Lakshminarayanan, B., Pritzel, A. & Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. *Adv. Neural Inf. Process. Syst.* **30**, 6405–6416 (2017).
- Yang, Y., Lv, H. & Chen, N. A survey on ensemble learning under the era of deep learning. *Artif. Intell. Rev.* **56**, 5545–5589 (2022).

37. Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning* (MIT Press, 2006).
38. Hutter, F., Hoos, H. H. & Leyton-Brown, K. Sequential model-based optimization for general algorithm configuration. In *Proc. International Conference on Learning and Intelligent Optimization*, 507–523 (Springer, 2011).
39. Bergstra, J., Bardenet, R., Bengio, Y. & Kégl, B. Algorithms for hyperparameter optimization. In *Advances in Neural Information Processing Systems*, Vol. 24 (Curran Associates, Inc., 2011).
40. Auer, P. Using confidence bounds for exploitation-exploration trade-off. *J. Mach. Learn. Res.* **3**, 397–422 (2002).
41. Iwasaki, Y. Dataset of B2-structured ternary alloys for magnetic moment, Curie temperature, and spin polarization calculated by KKR-CPA (AkaiKKR). *Mater. Data Repository Natl. Institute Mater. Sci.* <https://doi.org/10.48505/nims.5364> (2025).
42. López-Domínguez, V. et al. A simple vibrating sample magnetometer for macroscopic samples. *Rev. Sci. Instrum.* **89**, 034707 (2018).
43. Paulsen, M. et al. An ultra-low field SQUID magnetometer for measuring antiferromagnetic and weakly remanent magnetic materials at low temperatures. *Rev. Sci. Instrum.* **94**, 103904 (2023).
44. Fabian, K., Shcherbakov, V. P. & McEnroe, S. A. Measuring the Curie temperature. *Geochem. Geophys. Geosyst.* **14**, 947–961 (2013).
45. Soulen, R. J. et al. Andreev reflection: a new means to determine the spin polarization of ferromagnetic materials. *J. Appl. Phys.* **85**, 4589–4591 (1999).
46. Idzuchi, H., Fukuma, Y., Wang, L. & Otani, Y. Spin diffusion characteristics in magnesium nanowires. *Appl. Phys. Express* **3**, 063002 (2010).
47. Yaji, K. et al. High-resolution three-dimensional spin- and angle-resolved photoelectron spectrometer using vacuum ultraviolet laser light. *Rev. Instrum.* **87**, 053111 (2016).
48. Liang, Q. et al. Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains. *npj Comput. Mater.* **7**, 188 (2021).
49. Ngyyen T. D. et al. Stable Bayesian Optimization. In *Proc. Advances in Knowledge Discovery and Data Mining: 21st Pacific-Asia Conference, PAKDD 2017, Jeju, South Korea, May 23–26, 2017, Proceedings, Part II 10235*, 101–113 (Springer, 2017).
50. Klambauer, G., Unterthiner, T., Mayr, A. & Hochreiter, S. Self-normalizing neural networks. *Adv. Neural. Inf. Process. Syst.* **30**, 972–981 (2017).
51. LeCun, Y., Bottou, L., Orr, G. B.–Müller, K. R. Efficient backprop. In: Montavon G., Orr G. B., Müller K.-R. *Neural networks: Tricks of the trade* 9–50 (Springer, 2002)

## Acknowledgements

We thank K. Sodeyama, J. Hwang, and Y. Sakuraba at the National Institute for Materials Science and M. Kotsugi at the Tokyo University of Science for their valuable discussions. This study was supported by JST-CREST (Grant No. JPMJCR21O1). This study was supported by JST-CREST under Grant No. JPMJCR21O1.

## Author contributions

Naoki Yoshida and Yutaro Iwabuchi developed the code, analyzed the data, and contributed to the discussion. Yasuhiko Igarashi participated in the discussions. Yuma Iwasaki drafted the manuscript, designed the study, and managed the research.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-025-01851-8>.

**Correspondence** and requests for materials should be addressed to Yuma Iwasaki.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025