Designing workflows for materials characterization ⬤✓

Sergei V. Kalinin ✉ ⓘ ; Maxim Ziatdinov ⓘ ; Mahshid Ahmadi ⓘ ; Ayana Ghosh ⓘ ; Kevin Roccapriore ⓘ ; Yongtao Liu ⓘ ; Rama K. Vasudevan

Check for updates

View Online   Export Citation

15 April 2024 16:13:56

Applied Physics Reviews

AIP Publishing

Designing workflows for materials characterization ⒡

View Online · Export Citation · CrossMark

Sergei V. Kalinin,[1,2,a)] ⓘ Maxim Ziatdinov,[2,b)] ⓘ Mahshid Ahmadi,[1] ⓘ Ayana Ghosh,[3] ⓘ Kevin Roccapriore,[4] ⓘ Yongtao Liu,[4] ⓘ and Rama K. Vasudevan[4,c)]

AFFILIATIONS

[1]Institute for Advanced Materials and Manufacturing, Department of Materials Science and Engineering, University of Tennessee, Knoxville, Knoxville, Tennessee 37996, USA
[2]Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, USA
[3]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA
[4]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA

a)Author to whom correspondence should be addressed: sergei2@utk.edu
b)Electronic mail: maxim.ziatdinov@pnnl.gov
c)Electronic mail: vasudevanrk@ornl.gov

ABSTRACT

Experimental science is enabled by the combination of synthesis, imaging, and functional characterization organized into evolving discovery loop. Synthesis of new material is typically followed by a set of characterization steps aiming to provide feedback for optimization or discover fundamental mechanisms. However, the sequence of synthesis and characterization methods and their interpretation, or research workflow, has traditionally been driven by human intuition and is highly domain specific. Here, we explore concepts of scientific workflows that emerge at the interface between theory, characterization, and imaging. We discuss the criteria by which these workflows can be constructed for special cases of multiresolution structural imaging and functional characterization, as a part of more general material synthesis workflows. Some considerations for theory–experiment workflows are provided. We further pose that the emergence of user facilities and cloud labs disrupts the classical progression from ideation, orchestration, and execution stages of workflow development. To accelerate this transition, we propose the framework for workflow design, including universal hyperlanguages describing laboratory operation, ontological domain matching, reward functions and their integration between domains, and policy development for workflow optimization. These tools will enable knowledge-based workflow optimization; enable lateral instrumental networks, sequential and parallel orchestration of characterization between dissimilar facilities; and empower distributed research.

I. INTRODUCTION

Scientific progress is inherently linked to the development and utilization of progressively more complex methods for synthesis, imaging, and functional characterization of materials, from simple human eye-based examination and macroscopic property measurements to bespoke electron[1] and scanning probe microscopes (SPMs),[2] scattering facilities,[3–5] and low-temperature quantum measurements.[6–8] These imaging and characterization techniques, in turn, provide feedback for material synthesis optimization,[7] enable refining of theoretical models,[9] and often lead to serendipitous discoveries.[10,11] The role of tool development in science is reflected by the renowned quote by Freeman Dyson, one of the leading physicists of the 20th century: "*New directions in science are launched by new tools much more often than by new concepts. The effect of a concept-driven revolution is to explain old things in new ways. The effect of a tool-driven revolution is to discover new things that have to be explained.*"[12]

Present-day material discovery, design, and optimization are based on a well-established centuries-old paradigm of serendipitous findings of materials with useful functionalities, and long and time-consuming sequential optimization of compositions and processing conditions toward target functionalities. However, this approach tends to be extremely inefficient in the systems with multiple functionalities that achieve their optimal properties in different parts of multicomponent phase diagrams or synthesis parameter space. One of the material systems where these limitations are particularly important is the hybrid perovskite for solar cells and other optoelectronic applications.[13–17]

Similar challenges emerge for other multicomponent materials and devices including Li-ion batteries,[18] metallurgy,[19] high-entropy alloys,[20] and many functional ceramics and glasses.[21]

The material synthesis and characterization workflows typically emerge within specific domain areas, such as the epitaxial thin film growth community,[22] Li-ion batteries,[23] hybrid perovskite solar cells,[16,24] crystal growth in condensed matter physics[25] or radiation detectors,[26] and many others. For many of these fields, these workflows are inherently familiar to any practitioner in the field and often define it. Once novel or improved characterization tools appear, the workflows adapt to balance the availability of new tools, perceived gains in knowledge, potential for discovery, and costs in terms of time and expenses. This balancing is almost invariably based on intuitive decision making and is constrained by the availability of characterization tools, expected waiting times, and costs.

The new opportunity in the experimental domains is the rise of automated experiments (AEs),[27,28] where the artificial intelligence/ machine learning (AI/ML) methods are used both to enable automation to reduce latency (within a domain) and guide the discovery workflow. The combination of these two concepts gives rise to the concept of automated laboratories for the discovery of new materials for pharmaceutical and biological science and energy applications including solar cells. Despite some early demonstrations, this concept became mainstream only in the last 5 years, as a result of the large-scale efforts by Cronin *et al.*,[29–31] Maruyama *et al.*,[27,32–34] Aspuru-Guzik *et al.*,[33,35,36] Abolhasani *et al.*,[37–39] and others[17,40–56] as well as (less advertised) efforts in big pharma industries. For the last 3 years, the effort in small-scale laboratory-based automated experimentation (AE) via solution synthesis robot resulted in advancement in high-throughput and combinatorial studies of hybrid perovskite materials by Ahmadi *et al.*,[57–60] Brabec *et al.*,[61,62] etc. However, simple acceleration of material synthesis by typically reported ∼10 and even 2–4 orders of magnitude[63–65] for individual steps is insufficient compared to the vastness of the composition and processing spaces of multicomponent materials, necessitating development of workflows that will efficiently guide the synthesis and experimental protocols based on the results of previous experiments and general domain knowledge.

As an additional consideration, emergence of new tools gives rise to new opportunities for workflow development. For scanning probe microscopy, examples of this include the development of single-molecule unfolding spectroscopy that has opened the pathway to study the kinetics and thermodynamics of single-molecule reactions using benchtop tools,[66,67] piezoresponse force, and electrochemical strain microscopies that have enabled quantitative studies of bias-induced phase transitions and electrochemical reactions at the single defect level,[68–70] and scanning tunneling microscopy for exploring quantum physics[71,72] and chemical reactions[73] on a single atom level and enabling atomic fabrication.[74] For electron beam methods, the examples will include cryo-electron microscopy[75] that enabled mapping of protein structures and hence accelerated drug discovery, and electron diffraction[76] that allowed acquisition diffraction data from very small crystals for crystal structure determination. Examples abound in other fields. It is also important to note that many imaging tools can be also use for material synthesis, including the dip-pen lithography[77] and electron beam atomic manipulation.[78–80] Similarly, local methods can be seamlessly combined with well-established combinatorial spread

libraries[81] to establish the characterization feedback and further integrated with laboratory robotics.[82]

However, until now these developments have been largely *ad hoc*. The workflows have been developed in individual fields[83,84] and grow and evolve as a result of multi-year community-wide processes. New techniques give rise to fundamentally new scientific opportunities with often rapid growth, but discovery of these opportunities is often a black swan event rather than the result of long-term community-wide planning. Most importantly, in the everyday activity of research groups across academia, government labs, and industry, the choice of the measurements and characterization tools is determined by tradition far more then planning or analysis of possible gains and costs.

The workflow development can be subdivided into several elements including ideation, orchestration, and implementation, as defined in Table I. Traditionally, all three elements are human based, and the progression of the scientific career starts with the implementation and progresses to the orchestration and ideation part. The rapid emergence of networks of scientific user facilities and cloud laboratories disrupts this progression, allowing the implementation of workflows via computerized orchestration agents for human and automated equipment. The development of automated labs spurred the growth of instrument-level drivers and enterprise-level software ecosystems allowing for orchestrating the operation of multiple tools and creation of software layers for data storage and analytics.[85–88] The key step now is developing systematic ways to design, implement, and build the characterization workflows and determine the gains in terms of materials and physical discovery.[83] Ideally, we want to determine the sequence of synthesis and measurements in an optimal way, balance the cost of the tools and required characterization times to the knowledge or other gain, and use this to develop characterization workflows. We further want to adapt the workflows to the emergence of new characterization and synthesis tools, and therefore estimate the potential benefits from their introduction.

Here, we describe the scientific workflows, analyze their ideation and optimization for the specific case of the characterization methods, and explore the combined workflows containing characterization and synthesis components. We explore some elements of the workflow building when synthesis, characterization, and theory components are present. Finally, we analyze how these components can be used to enable the next generation of scientific research, including orchestration of the geographically distributed synchronous multimodal characterization workflows, lateral instrument networks, and the emergence of distributed experimental workflows across enterprise-level and community-level facilities. To accelerate the adoption of these tools by the community, we propose the framework for the workflow design, including the development of appropriate hyperlanguage and ontological connections between the domains, identification of a specific hierarchical reward functions, and policy development.

II. CLASSICAL SCIENTIFIC WORKFLOWS: IDEATION, ORCHESTRATION, AND IMPLEMENTATION

To illustrate the general concept of a scientific workflow, and some of the general principles of scientific workflow design, here we discuss several examples from areas that the authors are familiar with. However, similar elements and constructs can be identified across multiple other domain areas.

Shown in Fig. 1 is the example of the workflow development for the hybrid perovskite synthesis in the Ahmadi lab. Here, shown is the

TABLE I. Definition of workflow elements.

|	
Workflow	The sequence of steps, processes, or operations performed by human or ML agent when working in the lab or running the microscope to complete a specific task or achieve a particular goal. It outlines the sequence of tasks, their order, and the inputs and outputs of each task
Ideation	The process of generating and developing new ideas and approaches for workflow design and optimization. The goal of workflow ideation is to produce innovative and effective workflow designs that can achieve target goal (reward) more effectively
Orchestration	Management and coordination of a series of interconnected tasks or processes defined within the workflow that needs to be executed in a specific order to achieve a particular goal or outcome. It involves the automation and monitoring of workflows, ensuring that each task is completed correctly, and that the overall workflow is running smoothly
Execution	Performing specific steps or tasks in the workflow
Hyperlanguage	The language expressing operations that can be performed in the lab or running the microscope that define actions to be taken and their parameters. The hyperlanguage for ML agent needs to be expressed via API controlling the instrument
Reward	The perceived goal of the experiment. This can be discovery of material with specific functionality, optimization, or deriving general knowledge. The workflow is designed to maximize the reward. Rewards in scientific workflows are often hierarchical in nature.
Policy	The set of rules that define the actions taken depending on the state of the system. The important aspect of policies is the balance between exploration and exploitation, i.e., exploring new regions of parameter space or maximizing specific reward.
Value	As adapted from reinforcement learning, expected long-term reward or return that an agent can achieve by taking a particular action in a given state of the environment. The value of an action is calculated based on the expected sum of future rewards that an agent can receive from that action and all subsequent actions that follow.
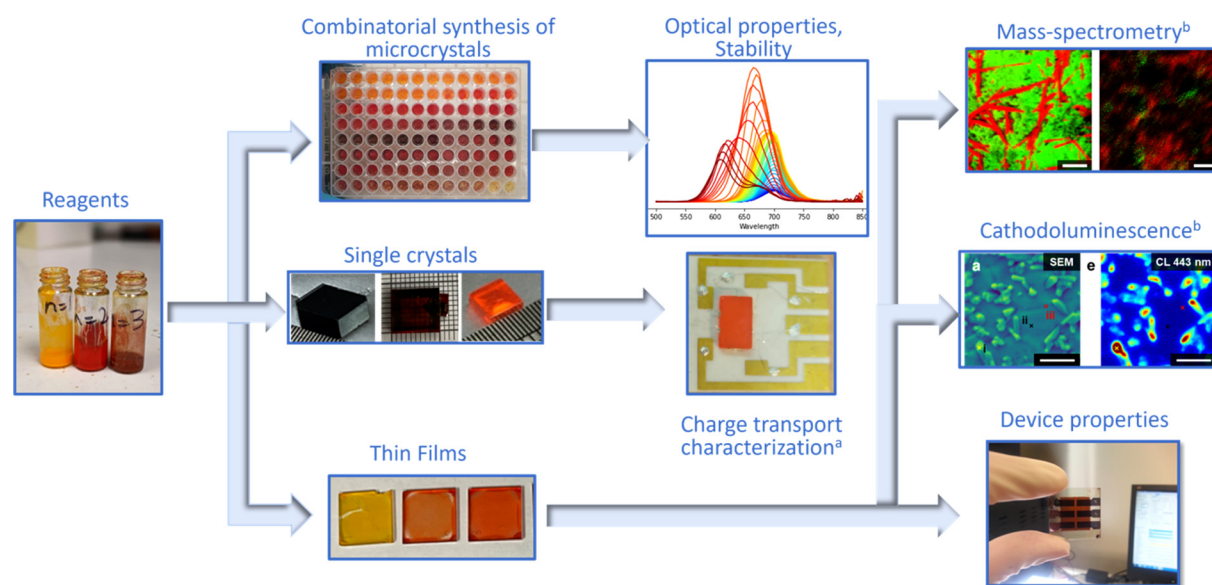
FIG. 1. Workflow design in the Ahmadi lab focusing on the hybrid perovskite synthesis. Similar reagents are used for the high-throughput synthesis and rapid photoluminescent emission assessment of bandgap energy, stability and quantum yield, single-crystal growth, and thin film deposition. Depending on the final form, the materials can be used to make devices and subsequent physical testing or explored by spatially resolved imaging techniques including cathodoluminescence and chemical imaging. The active devices, in turn, can be characterized by spatially resolved versions of these methods, providing insight into electronic and ionic dynamics. This forms the backbone of the operational workflow in the lab. The workflow ideation is based on multiple feedback steps where the results of the measurements of properties and functionalities inform the decision making for selecting compositions and processing conditions, stability measurements inform sample selection for film deposition and device fabrication, etc. Most of these decisions are made by human agent and are driven implicitly by combination of intuition on perceived reward, latency of measurement orchestration, and cost. (a) Reproduced with permission from Musiienko et al., Energy Environ. Sci. **12**, 1413–1425 (2019),[89] Copyright 2019 The Royal Society of Chemistry. (b) Reproduced with permission from Yang et al., Adv. Energy Mater. **13**(33), 2202880 (2022).[90] Copyright 2022 John Wiley and Sons.

material skeleton of the workflow, meaning the pathway the materials follow during the experiment. Note that workflow can also be defined in terms of the human or instrument time, depending on the specific problem to be solved. The initial points of the workflow development are the initial solutions and antisolvents. These can be used in a pipetting robot to prepare combinatorial libraries spanning the composition space of the hybrid perovskite, antisolvent compositions, etc. The optical bandgap energy, material stability, and quantum yield within the libraries are, in turn, accessed by the time-evolution of photoluminescent and UV-Visible absorption spectroscopies. Based on the results, the specific composition can be chosen for imaging studies with high spatial resolution. Similar solutions can be used for the single-crystal growth. The crystals can be fabricated into the devices for radiation sensors, explored via photo-Hall effect spectroscopy, neutron scattering, or Mossbauer spectroscopy. The crystal-based devices can, in turn, be visualized in operando via scanning probe microscopy or time-of-flight mass spectrometry (ToF-SIMS). The same initial reagents can be used to deposit thin films. The films, in turn, can be further fabricated into the device structures that can undergo physical testing. The microstructure of the films can be explored using cathodoluminescent (CL) imaging, ToF-SIMs, and scanning probe microscopy. Finally, devices can be characterized *in operando* using same techniques as crystals.

The characteristic aspect of this workflow is the progression from rapid, low cost, and high-throughput methods that provide limited and low-fidelity information on specific functionalities, to the expensive slow characterization methods for end materials or devices. In this process, the orchestration of the workflow includes multiple decision making and feedback steps on what composition to choose for complex characterizations. Note that the decision making is often nonlinear and gives rise to multiple feedback cycles at different latencies and degrees of analysis. For example, the results of the photoluminescence (PL) screening can be used for the composition selection for film deposition, and the CL and ToF-SIMS imaging of films will be used for composition selection for the robotic synthesis or the material selection for the initial endmembers or solvents. This decision making is further informed by the general information available to the researchers and adjusted as a result of the interaction with the scientific community via publications, conferences, social networks, interaction with large language model (LLM) optimized for generating hypotheses (a scientific version of ChatGPT), and private communications. We note that very recently the first systems allowing for such development have been introduced.[91–93]

The second example shown in Fig. 2 is workflow development in imaging.[83] Note that this workflow is hierarchical[94] element of the workflow in Fig. 2 and can apply to the cathodoluminescence,
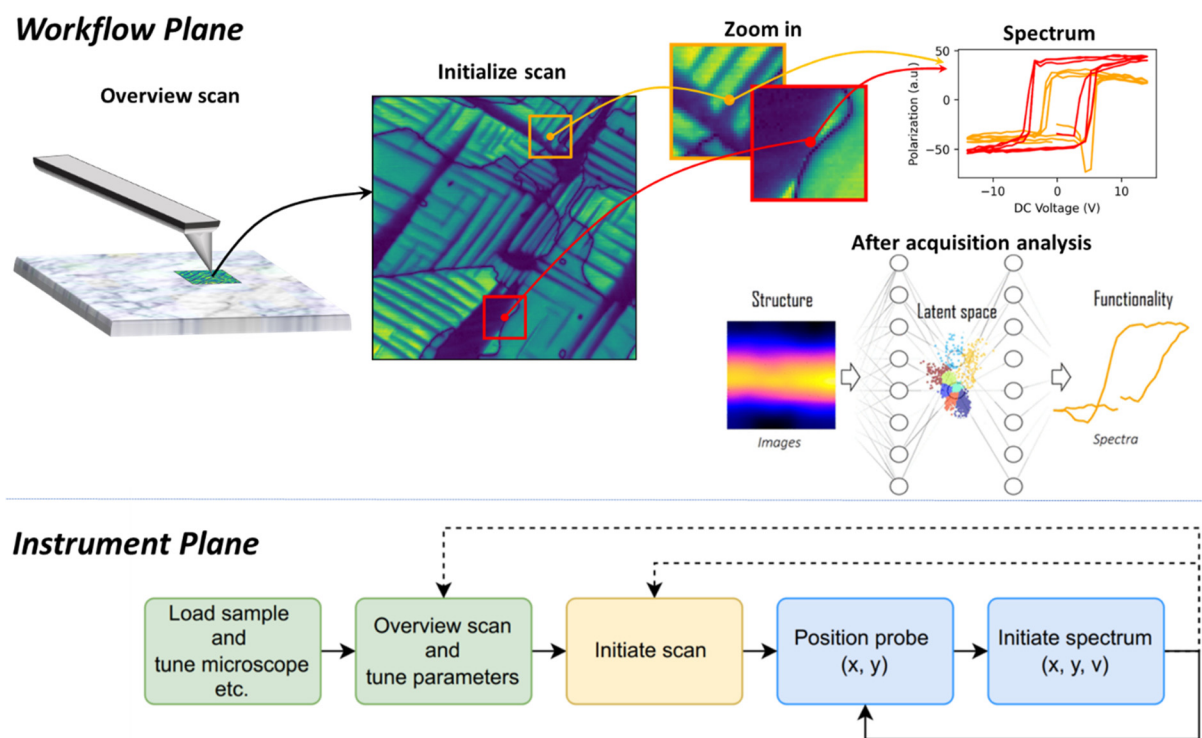
FIG. 2. Workflow development in imaging. Here, the decision making includes the selection of specific regions for high-resolution studies and subsequently for the spectroscopic probing. This process is often iterative and includes multiple overview scans, zoom-in and zoom-out stages, and human-driven spectrum acquisition and hyperspectral imaging. Note that for many instruments, the workflow will also include the repetitive tuning of instrument parameters for maximizing instrument performance. In this case, the workflow skeleton represents that the sequence of operations performed by the microscope, and workflow orchestration is effectively a stochastic optimization process. The nature of the reward function and values of individual steps are dependent on human operator.

ToF-SIMS, or scanning probe microscopy. Here, the natural backbone of the process is the series of operations performed by the microscope based on the input from human operator and internal feedback. The initial state of the system comprises the chosen sample and operator knowledge. The instrument operation includes tuning of the imaging conditions and overview scanning. Based on observations, the operator makes a decision to zoom-in on specific regions to explore these in detail, and potentially zoom-out and zoom-in on a different region and perform spectroscopic measurements. We note that for the unknown systems (for which there are no prior measurements or sufficient literature data), the selection of zoom-in regions is usually based on the educated guess of an operator and what he/she considers to "look interesting." Depending on the type of the measurement and instrument configuration, the operator can perform spectroscopic measurements on the grid (hyperspectral imaging). The process can also include the additional microscope tuning steps, with the human operator selecting the times to introduce them based on the operator assessment of the observed data.

There are three pertinent elements of the described process. The first is a clearly defined hyperlanguage that describes the human-initiated high-level operations executed sequentially (Table I). This language can differ between domain areas, but for most human-operated tools the set of elementary commands are similar. The second is the workflow orchestration based on the responses generated during the experiment and evaluated by human operator, including the decisions of what region to select for scanning and spectroscopy, and when to tune the instrument. Note that instrument automation makes some of these operations automatic, and often prompts the operator to perform them. The third and key element is the (scalar or multi-objective) reward function. The operator interested in mechanical properties of the material will be interested in different objects than the person exploring the emergence of ultrahigh electromechanical responses. Similarly, scientists interested in ferroelastic phenomena will choose different objects for study than those interested in flexoelectric phenomena at ferroelectric domain walls. This reward function, in turn, determines the value of the individual steps for the operator and, in this fashion, guides the workflow ideation. Finally, note that the scientific workflows and rewards have a clear hierarchical character, obvious given that example in Fig. 2 represents any of the imaging techniques that are used as a part of workflow in Fig. 1.

The brief examination of the workflows for synthesis and characterization illustrates several common elements. The first is the emergence of the funnel character. All samples are explored using easy characterization methods such as optical microscopy (and sometimes just human eye) or low-resolution overview scans, and this information is used to select locations for progressively more complex or expensive methods for smaller number of samples or locations. The second and less obvious component is the perceived reward of the experiment. In some cases, this is purely curiosity-based selection. In others, this is targeted exploration of specific aspects of microstructure or material behavior. In this latter case, the exploration pathway is driven by the specific interest of the experimentalist. The third implicit component is reliance on prior knowledge in selection of objects of interest, interpretation, and establishing the reward. Similarly, the discoveries can update the knowledge base, while slow for characterization experiments, this strongly affects material synthesis workflows shown in Fig. 1.

The two examples above illustrate the workflow concept, organization of the workflows based on materials or instrument time, and the complex hierarchical connections between the reward functions of the individual elements and values of the specific steps and operations. Below, we explore the principles based on which these workflows can be designed and optimized—here, for the specific case of characterization workflows as shown in Fig. 2. However, we emphasize that these workflows are, in turn, defined only in the context of general synthesis and characterization workflows shown in Fig. 1. In other words, materials first need to be synthesized—and the motivation for synthesis is derived from combination of specific goal (material for photovoltaics) and discovery. Characterization and microscopy aim to provide the feedback to material synthesis workflow.

In discussing these workflows, we separate the components of imaging, spectroscopy, and theory. Here, imaging is referred to the process of the acquisition of spatially resolved structural and functional information. Spectroscopy refers to the process of the detailed measurement in a specific location that is assumed to provide desired information on material functionality. Note that the process is hierarchical, in a sense that imaging can be performed via the acquisition of spectra on the sample grid (hyperspectral imaging) if the spectroscopy is nondestructive. Due to the difference in cost and latencies, this gives rise to such standard tasks as pan-sharpening.[95] Finally, the third component is theoretical analysis, generally referring to the derivation of the insight from observations and using these to modify the way workflow is ideated. First, we discuss the principles of workflow development for pure imaging and spectroscopy scenarios and compare it to the (well explored) ideation of theoretical workflows.

III. HOMOGENEOUS WORKFLOWS

We define the homogeneous (or single element) multiresolution workflow as those emerging within a single type of characterization or theory connecting multiple length scales.[96,97] Here, we consider the well-known multiscale workflows in theory and expand these concepts to characterization.

A. Multiscale workflows in theory

The concept of multiscale workflows is well developed in the theory domain. Over the years, the development and implementation of multiscale modeling[98–105] workflows have paved their way into multiple disciplines ranging from solid mechanics, fluid mechanics, materials science, physics, mathematics, and biological to chemistry. Parallel computing has only made it more feasible to solve more accurate and precise algorithmic formulations, which is needed for these workflows. In addition to being useful in academic research, such modeling capabilities have been adapted in industry for its numerous advantages such as cost-effective physics-based product design, assessment of product quality, and performance. The primary requirement for any multiscale workflow is to lay out strategies to bridge between different length scales, going from atoms to automotives. We note that it differs from the conventional point of view being followed in various disciplines where the focus remains on solving a particular challenge with sole consideration of the pertinent length scale and associated latencies.

Two of the most common approaches as followed in the multiscale paradigm are concurrent[106,107] and hierarchical[94,108] in nature. Methods to bridge between length scales vary between the two. In

concurrent techniques, the bridging methods depend on numerical solutions, whereas hierarchical approaches rely on performing independent numerical simulations at different length scales followed by identifying relations between parameters relevant to integrate or reconstruct material behavior at the corresponding higher layer in the ladder of length scales. The hierarchical approach is top-down in nature. Understanding microstructure–property relationships inferred from the interplay of internal state variables existent in various scales with thermodynamic constraints is a good illustration of such an approach. One example where concurrent and hierarchical schemes are practiced together[101,109–111] is in connecting atomic scale simulations with electronics principle scale simulations. Density functional theory (DFT) simulations performed on metals are coupled with embedded atom method (EAM) potentials within molecular dynamics (MD) environment to model edge dislocations, with subsequent studies performed with quantum models at disparate length scales. Here, the many-body interactions are evaluated within the semiempirical formalisms of the potentials such that the results from the electronic structure theory computations become useful to reproduce physical properties of many metals, defects, and impurities. In addition, it is also possible to conduct such multiscale studies on-the-fly where the classical potential adapts to the local environment via dynamic force matching. Several machine learning frameworks have also proven to be useful to develop these potentials[112–118] within an interactive suit bridging between the atomic, coarse-grained descriptions with promises of connecting to the continuum theories.

Machine learning methods have significantly accelerated the development of multiscale workflows for theory. Multiple approaches to address it have been used, ranging from rare event samplings in molecular dynamics (MD) simulations to machine learning-based information compression schemes.[119,120] Particularly over the last five years, a number of machine learning approaches based on variational autoencoders,[121] generative adversarial networks, and diffusion models have been suggested to bridge length and time scales in simulations, establish statistically significant descriptors such as order parameters, and determine their constitutive relations. It should be noted that many of these methods have also been shown for the information compression in experimental data such as electron and scanning probe microscopy,[122–124] albeit with the additional requirements to account for the out-of-distribution shifts due to changes in imaging conditions that are typically absent in modeling Fig. 3.

B. Multiresolution characterization workflows

The nature of the multiscale problem in experiment is opposite in the sense of information flow. Multiscale characterization in experiment requires solution of the inverse problem, namely, the optimal hierarchical experiment design. For example, given the results of macroscopic measurements or low-resolution imaging, we seek to identify the potential object of interest to be explored via high-resolution imaging probes in such way as to gain the maximal insight into the nature and origins of observed macroscopic behaviors.

To set this problem in a more mathematical basis, we consider it as one of optimizing a sequence of actions to maximize a cumulative reward function (Fig. 4).[125] The reward function can be defined in multiple contexts from pure optimal structure discovery (we aim to characterize the structure as detailed as possible via reduced multiscale representations) to discovery based on prior knowledge (we know what microstructural objects we are interested in). In the case of
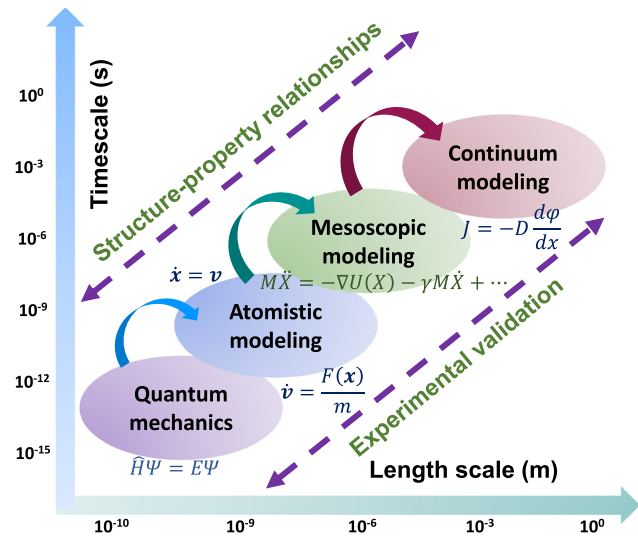

FIG. 3. Multiscale workflow depicting the integration of data from diverse time and length scale simulations, accompanied by a representation of the underlying cause–effect relations between structural parameters and material properties.

information gain, one can simply seek to minimize uncertainty arising from models at different length scales, by finding the sequence of measurements that can best reduce the overall uncertainty across length scales. Alternatively, a measurement sequence can be found that minimizes the uncertainty at one desired length scale. To construct a reward function for this can be as simple as taking the negative of the sum of the uncertainty. In a reinforcement learning (RL) framework, the actions consist of both the length scale to explore next in the workflow, as well as parameters within the specific experiment. For simplicity, we will ignore the latter component and consider that the only action is to choose the appropriate length scale. The task is to find a policy that will determine how best to select actions to maximize the cumulative rewards. Additionally, it is known that measurements at one length scale can be highly informative of measurements at other length scales.

For this scenario, the knowledge discovery process can be represented via a probabilistic machine learning (PML) framework based, for example, on Gaussian process, Bayesian neural network, or deep kernel learning (DKL), (Fig. 5), with the idea that the data can be mapped from one length scale to another. Assume that we have training data X captured at the nth level, X^n. We can generate predictions on the test data X_*^n,

$$X_*^n \rightarrow \overline{f}_*^n, \mathbb{V}\left(X_*^n\right). \tag{1}$$

As an example, we consider that we take a few optical images at certain positions (X^n), and then want to predict the image at unseen locations (X_*^n). The predicted images are $\overline{f}_*^n$, and the predictive uncertainty is given by $\mathbb{V}(X_*^n)$. However, we can also setup models that attempt to predict at the next (e.g., lower) length scale. By feeding in the test data X_*^1, we can train a PML model to predict the function value at, e.g., the next level, i.e., $X_*^1 \xrightarrow{PML^{(1)}} \overline{f}_{X_*^1}^2, \mathbb{V}^2(X_*^1)$. In general, we have

$$X_*^m \xrightarrow{PML^{(m)}} \overline{f}_{X_*^m}^k, \mathbb{V}^k\left(X_*^m\right), \tag{2}$$
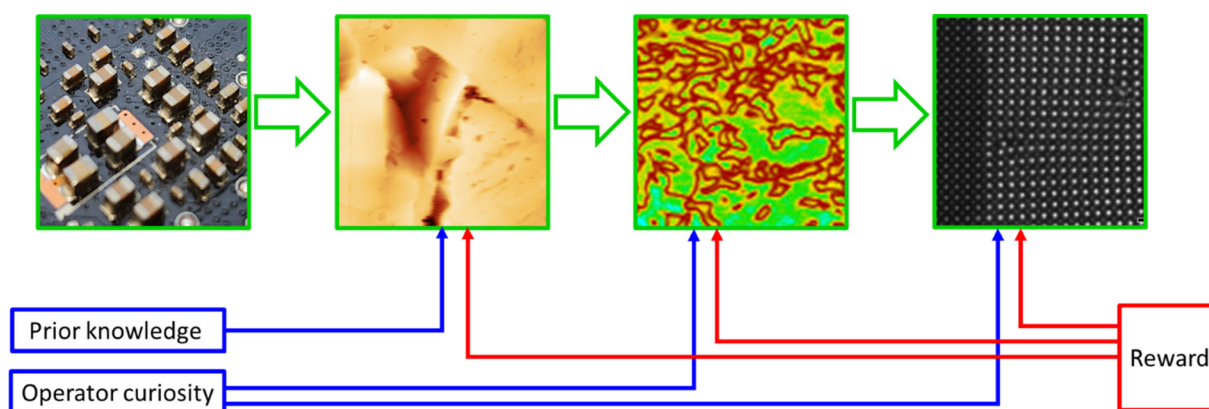
FIG. 4. The workflow for structural imaging. Here, the imaging studies are performed increasing the resolution of the microscope (and changing the imaging system). The decision making along the workflow includes selection of regions for detailed studies at progressively higher level of details. The decision making along the workflow can be purely data driven (green lines), or incorporate decisions made with prior knowledge and informed by perceived reward.
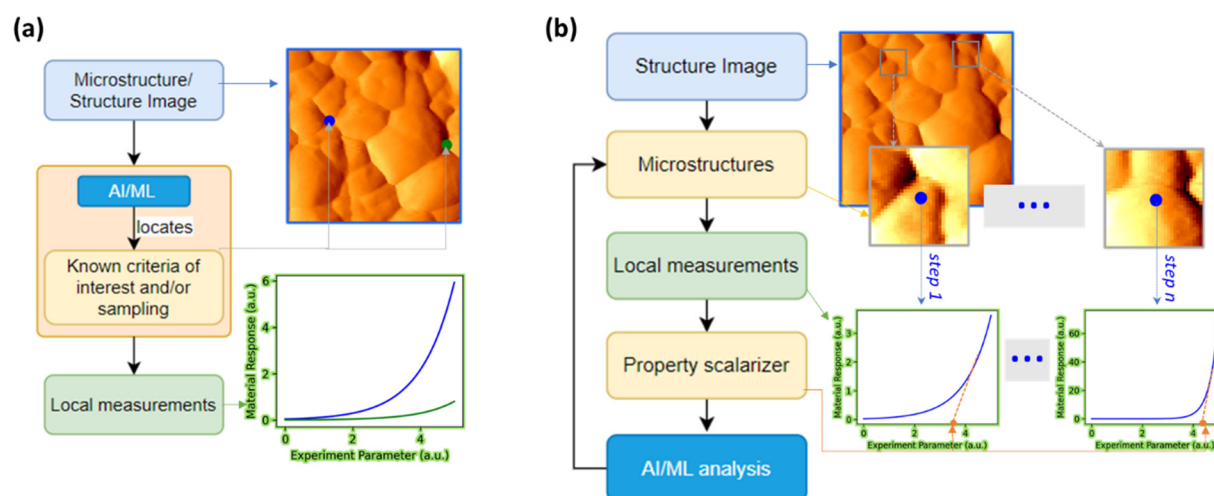

FIG. 5. Single-step workflows based on (a) known criteria of interest and (b) active exploration, also known as (a) forward and (b) inverse experiments. In the case of (a), the object of interest is known *a priori* and the role of AI algorithm is to identify them and perform specific measurements. In case (b), the scientist is interested in certain aspects of the functional properties defined via scalarizer, and the AI algorithms learn what microstructural elements maximize or minimize them.

where the measurement at level m provides estimates of the function value at level k along with uncertainty estimates at level k. The task is to determine which measurement level m (out of the n available levels) to choose that will minimize the uncertainty in $\mathbb{V}(\cdot)$, which could be the uncertainty at one level, at all the levels, or at some combination of levels, e.g., with a weighted sum. This means we need to know which $\mathbb{V}^k$ or combinations thereof to use. Minimizing this uncertainty will be the objective of the policy. After fixing the reward function, the workflow reduces to a simple reinforcement learning (RL) environment in which the action is to decide which measurement (level) to capture, which will then enable the appropriate $PML^{(m)}$ model to be updated. The state fed back to the agent will then be the set of predicted means, i.e., $\bar{f}_{X_*}^k$ for $1 \leq k \leq n$. In the case where we do not know which $\mathbb{V}^k$ to use, this can be reformulated as an RL problem, with the difference being that the action now selects not only the measurement level, but

also, which uncertainty map level is used. Alternatively, one may also consider some type of superpositions of the uncertainty from different levels. Additionally, given that different experiments will likely greatly differ in their actual cost per data point measured, this can also be factored into the reward function: penalties can be applied to actions or action sequences that use "expensive" characterization tools, for instance. Regardless, the optimization is straightforward once the reward is defined, similar to our recent work with hypothesis learning.[46,126]

As a practical example, we consider that we might have optical images, SEM images with some chemical maps (e.g., from energy dispersive spectroscopy, EDS), and some scanning tunneling microscopy images of a material system, such as a substrate with 2D flakes of varying chemical composition. This defines three levels, and the question is which set of experiments will be most informative. In this case, optical

microscopy of similar looking flakes is not likely to be highly correlated with local electronic structure as imaged by STM, if these flakes are different in chemical composition and defect concentrations. On the other hand, there will be a significant correlation with the EDS data on the same flakes. As such, after initial correlations between STM, optical, and SEM/EDS data are found, minimizing total uncertainty may hinge more on performing a few experiments with SEM/EDS and confirming them with STM. Our problem is in determining exactly which of these measurements are most informative, and this can be done, in principle, via the PML framework described above.

C. Beyond data driven discovery

The recent advances in imaging and characterization tools including electron and scanning probe microscopy and associated spectroscopies, atom probe tomography, focused x-ray scattering, nanoindentation, optical microscopy and a gamut of electrical, mechanical, and magnetic testing methods span the multitude of length scales and functionalities. However, the gamut of available techniques is belied by the dearth of the systematic workflows that allow exploring the behaviors of interest in a systematic and unbiased way. It is by now common to complement the macroscopic probing of the piezoelectric, catalytic, and electric properties by the STEM or atomic probe tomography studies of atomic structures, or correlate the photovoltaic performance of polycrystalline solar cells by the nanometer resolution cathodoluminescence and chemical imaging maps.

However, it is seldom that we can certainly say what specific type of microstructural elements is most strongly associated with the functionalities of interest. In the cases where such relationship was defined, as for the role of step edges for catalysis, dislocation theory of plasticity, or role of the domain wall dynamics on giant electromechanical responses in piezoelectric ceramics, these discoveries required community-wide effort involving multiple experiment and theory development cycles. At the same time, it is these insights that are most relevant toward understanding the underpinning mechanisms and particularly establishing pathways for material optimization and discovery. Only by understanding the fundamental origins of structure–property relationship can the strategies for the improving material performance be formulated and tested.

For example, given the polycrystalline ferroelectric material, we may seek to understand the origins of the high electromechanical response or resistive switching. Given the hybrid perovskite solar cell, we want to understand which microstructural elements are responsible for chemical stability, current–voltage (IV) hysteresis, or the open circuit voltage (OCV) losses compared to the ideal values. The complexity of such analysis stems from the fact that it may be different defect populations that are ultimately responsible for these behaviors, and hence, selection of the regions of interest for the detailed studies depends on the specific goals. In this sense, this problem is poorly defined—we seek to preferentially explore via detailed high-resolution studies objects whose identity we do not know. Hence, we use hypotheses formed based on the prior body of knowledge to guide the exploration process, while maintaining the need for serendipitous discoveries.

D. Reward design

These considerations further bring us to the concept of reward engineering. Here, methods such as reinforcement learning (RL) have been shown to be highly effective in simulated environments such as computer games or simulations.[127–129] One of the key elements of RL is a reward function that is made available for the algorithm during the training. However, for many real-world problems, the reward functions available in the end of experimental campaign (or after several steps) are absent; rather, the experiments motivated by the long-term objectives. Designing reward function that adequately represents real-world objective and does not lead to reward hacking[130,131] is a challenge. Similarly, very often experimental results can contribute to multiple objectives, with fundamental scientific research being the most notable example of such activity.

As an example of such a problem, we consider climate change, the problem motivating multi-billion-dollar investments over the globe. Minimizing climate change is a very long-term objective. The lower rank objectives are the development of solar and wind energy and associated grid-level storage and effective energy transport methods, along with the technologies for direct carbon capture. The even lower rank objectives are the development of cheap, environmentally friendly, and stable chemistries for grid storage. None of these objectives can be translated into a reward for an experimental campaign. Rather, these objectives serve as a motivation for experiment planning—and reward is often a short-term battery performance or observation of specific mechanism in microscope that can suggest potential ways to improve the battery materials.

We pose that discovery of the short-term rewards that can be used for hypothesis making to guide experimental research, and as rewards functions to guide and ascertain the success of experimental campaigns is the missing link required to connect ML to real-world applications. The potential pathways to address this challenge can include literature mining toward building the DAGs connecting experimental outcomes (rewards) and objectives (motivation), technoeconomic analysis of past publications outcomes, and crowdsourcing to the community of experts (aka "what would be the potential of high temperature conductivity to change the world" to "how does the phase separation in cuprates affects peak effect and losses").

The key consideration for the reward-driven workflow design is the capability to separate the specific objective into the probabilistic graph of short-term reward functions that can guide experiment planning and establish measures of success. Naturally, these reward functions will be probabilistic, and the value of real-world experiment can affect (much) more than one objective. For example, mechanisms of metal–air interactions can be used both for corrosion mitigation and for design of metal–air batteries. The important element of this approach is that humans are the part of the theory-experiment loop—and hence, the structure of the rewards can be amended via human feedback the observations (much like science works now).

Notably, the LLMs such as ChatGPT are often capable of making the connection between high- and lower-level objectives (e.g., prompts "what should I study with microscope to understand plasticity" gives very plausible answers). Presumably, complementing LLMs with models trained on domain-specific literature can both allow systematic developments of such workflows and their integration across multiple domains following common rewards.

IV. MULTIRESOLUTION STRUCTURE–PROPERTY CHARACTERIZATION WORKFLOWS

It is well recognized that understanding structure–property relationships in materials requires exploring properties and functionalities

on all length scales, from the atomic scale structures to mesoscale to the global properties of the material or device. Multiresolution analysis workflows are used when, for each characterization method, we have access to image–spectroscopy pairs, and the spectroscopic data provides information that is predictive of the functionality of interest. Here, we define as the image the high spatial resolution/low information density imaging including structural STEM image, SPM images, etc. The spectroscopy refers to the information-rich local measurements that are associated with larger measurement times or lead to the irreversible changes in material structure, and therefore can be performed only in a limited number of locations. However, we implicitly assume that the spectroscopic measurements are correlated with the macroscopic functionality of interest.

A. Single-step direct workflow

Here, first we illustrate these concepts for the single structure–property mapping step. In this case, we can define *forward* and *inverse* experiment. The forward experiment relies on the *a priori* defined objects of interest that can be recognized in real time, e.g., using deep convolutional networks. Here, the emergence of the ensemble and iterative training methods allowed to partially address the inevitable out-of-distribution effects (i.e., capability of the trained network recognize object of interest if microscope parameters have changed). Recently, a deep residual learning framework with holistically nested edge detection (ResHedNet) was ensembled to minimize the out-of-distribution drift effects in real-time SPM measurement.[132] The ensembled ResHedNet was implemented on an operating SPM, where it converted the real-time SPM data stream to segmented objects of interest, e.g., ferroelastic domain wall or polycrystal grain boundary images. Then, a pre-defined workflow used the coordinates of the discovered objects for spectroscopic measurements. In doing so, the approach allowed a thorough of interested objects of interest (virtually all locations at objects of interest) in an automated manner; in contrast, traditional manual operation only allows us to investigate a limited number of locations at domain walls. Using this approach, alternating

high- and low-polarization dynamic ferroelastic domain walls in a $PbTiO_3$ thin film was observed[52] and the behavior of grain boundary junction points in metal halide perovskites was discovered.[133]

B. Single-step inverse workflow

In the inverse experiment, the operator defines the characteristics that make spectrum "interesting," e.g., intensity of a specific feature, specific aspect of spectrum shape, or even maximal variability of spectra within the image. In other words, each collected spectrum can be associated with a single number defining how interesting it is, either in absolute sense or as compared to previously acquired spectrum. The deep kernel learning (DKL) algorithm learns what elements of the material structure maximize this reward and guides the exploration of material surface accordingly. This DKL algorithm was recently implemented in SPM to investigate the relationship between ferroelectric domain structure and polarization dynamics,[54,134] and in STEM to explore bulk and edge plasmonic modes. As shown in Fig. 2, the DKL exploration process identifies the domain walls as objects of interest and the DKL predictions indicate the high polarization dynamic of $180°$ domain walls. Although these are expected by ferroelectric experts, the DKL itself did not have any prior physical knowledge and learned all that information during the experiment.

C. Multiple-step workflows

The multiple-step imaging and characterization workflows can be represented as the direct extension of single-step workflows as shown in Fig. 6. Here, the structural imaging at the low resolution yields the library of possible microstructural elements. We assume that the microscopic measurements, e.g., using the micropatterned contact arrays and current–voltage (IV) measurements via SPM, are representative of the macroscopic properties of the systems (even though the exact physical mechanisms can be different due to changes in contact conditions, confinement effects, etc.). These elements can be sampled in statistically balanced way, e.g., based on the distributions in the latent space of the system, to give the initial information on the
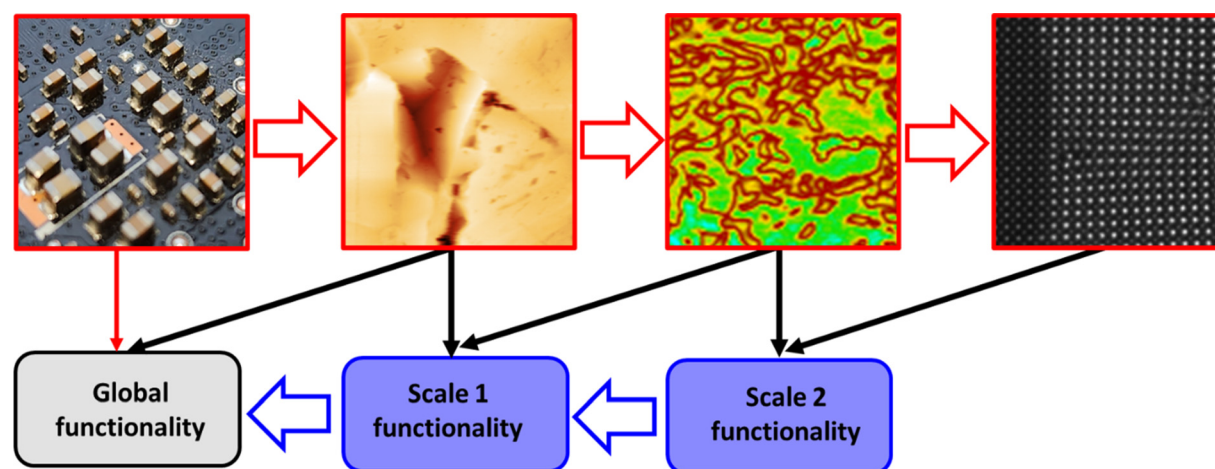

FIG. 6. Multiple-level imaging and characterization workflow. Here, we assume that the material functionality of interest probed on the macroscopic level is controlled by the hierarchy of structural elements from macroscopic to atomic scale and that the available functional probing methods are representative of the material functionality. For example, on the mesoscale, this can be realized using the microelectrodes and the SPM-based IV measurements.

structure–property relationships in the system. With these, the process can be iterated balancing the structural learning and learning structure–property relationships on multiple length scales. For the cases when the spectroscopic measurements provide information that is a proxy to the macroscopic functionalities but does not directly reproduce it, the definition of the workflows becomes more complex and requires incorporation of the multiresolution multifidelity measurements. For example, this can include the use of the easy to measure signals such as micro-Raman to identify preferential locations for expensive measurements such as scanning probe microscopy or scanning transmission electron microscopy, as has recently been explored by Kusne.[48]

The key emerging aspect[135–137] will be the integration of workflows across diverse domains, including different characterization methods and synthesis. This, in turn, necessitates the employment of consistent ontologies, a critical factor for ensuring effective interdomain communication and data interoperability. Within multidisciplinary landscapes, each field typically develops its own unique lexicon, data structures, and conceptual frameworks, which can lead to significant challenges in cross-domain interactions. The disparate nature of these elements often impedes the seamless exchange and accurate interpretation of information, posing a barrier to the cohesive integration of workflows. Consistent ontologies should offer a standardized, unified schema for representing and understanding diverse data sets and inferential biases, enabling disparate systems to interact synergistically. This standardization is imperative for the accurate mapping of concepts and terminologies across different fields, ensuring not just data exchange but also its meaningful contextualization and utilization. Furthermore, the adoption of consistent ontologies facilitates scalability and adaptability within integrated systems, allowing for the incorporation of emerging knowledge and technologies.

V. SOME CONSIDERATIONS ON THEORY–EXPERIMENT–CHARACTERIZATION WORKFLOWS

The particularly interesting problems emerge at the interface between the theory and experiment. The conflux of parallel computing, experimental capabilities along with theoretical simulations is necessary to develop and implement end-to-end theory–experiment–characterization workflows. It is safe to say that within materials science challenges spanning over various applications, the primary goal is to leverage existing structure–property relationships[138] to propel both design and discovery. These relations could be formulated at the atomistic level where electronic structure of systems plays the key role to determine the energetics and stability. For mesoscale to continuum scale, we tend to map the evolutions of microstructures with physical properties such as plasticity, damage, or failure. The standalone investigations[139–148] based on data from theoretical simulations[149–154] are well capable of elucidating the fundamental mechanisms responsible for materials characteristics with applications in energy, catalysis, and photovoltaics, drug design, to name a few.[118,155–171] However, the ultimate validations of such proposed mechanisms always rely on experimental observables. Hence, to fully realize the potential of workflows[161,165,168,172–176] bridging instruments and theory,[174,177–191] we must move toward theory-assisted experiments from the standard perspective of matching final outcomes from experiments and simulations. It becomes important to consider how causal structure–property relations[192–194] may hold true at different length scales while establishing connections between experimentally controllable parameters with

theoretical variables. Here, we consider this interaction only in the more limited context of imaging and characterization methods, but even in this case, the complexity of possible interactions is immense.

Here, we assume that we have access to microstructure (M) at a single length scale, global property measurements (G), local property measurement (L), and theoretical model (T). We further assume that M and G are available from the beginning of the experiment and are not updated. Comparatively, T is available in the form of analytical or numerical model and can contain partially unknown parameters that can be updated during the experiment, and L can be performed sequentially within known M. With this, we can define the static learning problem, meaning establishing the relationships between G, M, and T after the measurements, and active learning problems, meaning the workflow design for active experiments within M. Here, we will use arrow → to define establishing a relationship given full batch data, and ▶ to define active learning workflow.

A. Static problems

With these building blocks, we can combinatorically define three static analysis problems. Here, MT → G is a direct calculation problem. In this case, we assume that the microstructure is known and the theory is correct, and aim to calculate the global properties of the system. The example of such approach will be the finite element calculation of the mechanical properties of the composite material, estimation of the effective transport properties of the microstructure mapped via the x-ray tomography, etc.

The inverse problem will be MG → T. Here, given known microstructure and global properties, we aim to refine theory that governs structure–property relationships in the system. Finally, GT → M defines the design problem. Given the property of interest and theory, we aim to design microstructure satisfying the given properties. These problems are static in nature, i.e., we do not have iterative/active learning component. Of course, these problems become active when a part of the synthesis or manufacturing workflow. Similar workflows emerge in the context of local property measurements.

B. Active learning problems

With these simple examples, we can define several classes of active learning problems. As mentioned above, defining active learning workflows necessitates introduction of the reward, R, defining the discovery target. For example, M(R)▶L is the direct imaging problem, where we choose locations for sampling local microstructure L given what is interesting about M. In our experience for human-driven exploration, the reward function often changes during the experiment. For example, initial measurements are done with the target of microscope performance optimization, proceed to get the overview image, explore the most statistically prevalent regions within the image, and then proceed to explore the regions that are believed to be interesting based on the prior knowledge and beliefs. For example, it is natural to target dislocations for mechanical properties, or ferroelastic domain walls for understanding of the origins of the ultrahigh electromechanical responses.

Comparatively, the L(R) ▶ M is the DKL problem, where we aim to discover locations for L given what is interesting about L and perceived reward R. We can also envision more complex workflows, for example, ML ▶ T, meaning how to learn the theoretical model in

an efficient manner given microstructure and local measurements. The more complex scenarios include M G(R)▶L, M T(R) ▶ L, and M GT (R) ▶ L, meaning that we aim to discover locations for L given what we know about microstructure and theory and perceived benefit for global properties, theory, or both. As defined earlier, the reward can be optimization of some characteristic, or simply, information gain as represented by reduction in uncertainty of DKL (or other relevant) surrogate models.

For multilevel problems, actively learning theory at the same time will induce additional nonstationary characteristics to the objective, further increasing the complexity of the problem. In fact, even the simple case of the co-navigation of theoretical and experimental domain requires detailed analysis balancing the uncertainties between the domains.[195] Similar to single domain workflows, the development of consistent ontologies for description of objects (and even search spaces) is a key part of these developments.

VI. IMPLICATIONS

Finally, we analyze how the design and optimization of synthesis and characterization workflows can affect the structural organization of the research process. Specifically, we consider both the driving forces such as increase in the throughput and costs of many imaging and characterization tools, and emergence of the cloud infrastructure and edge computing that allow information processing and feedback from the cloud.

A. User facilities and lateral instrumental networks

From the historical perspective, the first big changes in the scientific workflow in condensed matter physics, materials science, and biology were brought about by the emergence of the large tools such as synchrotrons and nuclear reactors.[196,197] At that time, the concept of user facility included developing the instruments, instrument scientists operating them, and user scientists that physically visit for day-week long experiments on their specific materials has become a norm.

The second big change has been emergence of the user facilities, as exemplified by the Department of Energy Nanoscale Science Research Centers, and user facilities at universities. This trend is associated with the emergence of the microfabrication labs as a part of exploratory research, and rapid growth of throughput and costs of characterization tools such as electron microscopes, scanning probes, and chemical imaging. Correspondingly, integration of these tools within the same geographically localized facilities that maintain the instrument, offer the sample preparation facilities often shared between multiple instruments, and maintain the highly trained scientists capable of operating and using them greatly increases the efficiency of use.

However, despite the drastic changes on the operational side, the mode of use of the instruments in user facilities and individual labs has been remaining largely the same. In all cases, the scientist run the instrument manually, generating the large volumes of data during the experiment. The data are typically analyzed after the experiment, in the process that often takes weeks and months, and is subsequently shared with the scientific community via publications. The latter process is typically associated with extremely large latencies of the order of months and often years, hindering the research process. Correspondingly, until 10–15 years ago, scientific conferences were the primary means for rapid information sharing. Over the last decade,

rapid growth of popularity of the preprint servers such as arxiv, chemrxiv, and biorxiv, as well as social networks greatly accelerated information sharing. Similarly, code and data sharing via platforms such as GitHub, Zenodo, and Google Colab are rapidly becoming a part of scientific culture in many domain areas.

The rapid increase in ease of use of cloud technologies suggests that the field is now poised to the next transition, where the operation of the tools is largely remote and the information is either directly streamed to the cloud storage, streamed after the acquisition based on the upload speed constraints, or following the point-of-generation compression. In addition to obvious advantages in terms of data accessibility and security, this allows the formation of the lateral instrument networks as illustrated in Fig. 7. Here, multiple instruments of the same kind store the data within a community-accessible cloud space. The latter also supports the computational capabilities and code ecosystems that allow data analysis and, in turn, can be further harnessed for the decision making and automation of workflow on individual instruments, significantly accelerating the scientific process. It should be noted that very likely data permissions will be dependent on the study; most likely, data will not be shared universally without an embargo period or something of the sort.

B. Sequential and parallel experiment orchestration

The development of cloud connectivity for characterization tools establishes a set of novel opportunities for experimental workflows across multiple facilities. The particularly important case for the parallel experiment orchestration[198] is when multiple copies of the same sample are available, e.g., combinatorial spread libraries. In these, the local concentrations and functionalities are rigidly encoded in positional descriptor, allowing matching the regions between different tools. For these structures, only a few characterization methods such as optical hyperspectral imaging or photoluminescence can be performed in parallel. For techniques such as structural characterization via focused x ray, chemical characterization via ToF-SIM, cathodoluminescence, and scanning probe microscopy, the measurements can be performed sequentially. At the same time, these techniques often give complementary information on structure, properties, and chemical composition. Correspondingly, running the automated experiment on same object and multiple systems allows to explore the material composition space combining information from multiple sources (Fig. 8).

We pose that these experiments can be also performed in the statistical sense, in which the information from multiple tools is combined via partially similar channels. For example, for hybrid perovskite samples in Fig. 1, the statistical properties of the microstructures can be explored via correlating the chemical or CL signals referenced to grain boundaries or other key point objects that can be identified in both methods.

C. Automated laboratories vs user facilities

It is important to distinguish between cloud laboratories and user facilities, as they serve different needs. The former is generally a replication of what would be found in an individual lab at a university, though with APIs to enable automated experimentation and data collection, and the possibility of enabling autonomous workflows, given the presence of these APIs. This means that individual researchers can
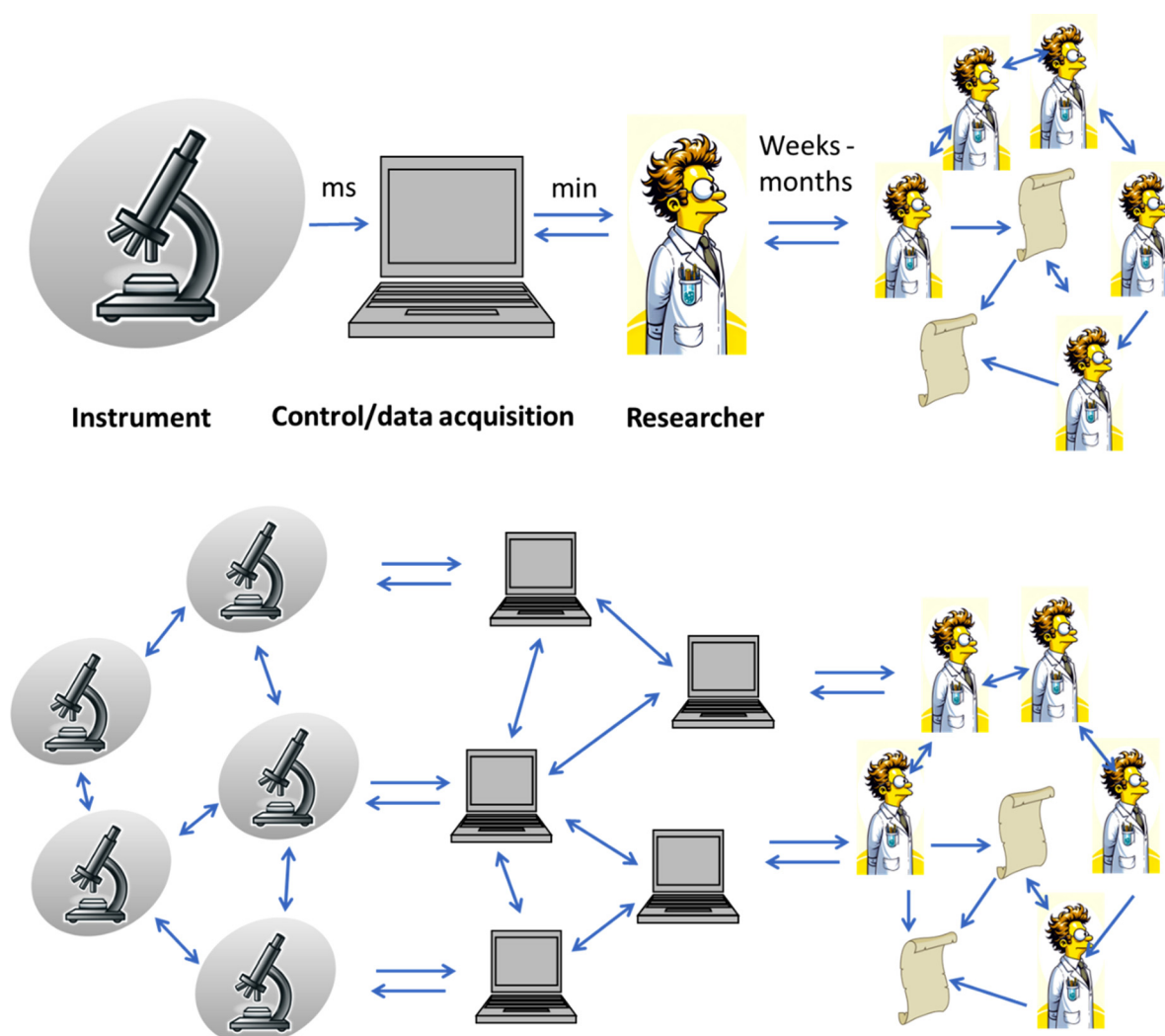
15 April 2024 16:13:56

FIG. 7. Transition from single tool to lateral instrumental networks enabled by cloud technologies and data analysis ecosystem. Traditional scientific research often relied on individual, isolated instruments. Each instrument was operated independently, and the interaction between different instruments was performed in the phase of post-experiment data analysis by human. The transition here is to move from this isolated setup to a network of instruments that are interconnected, which allows for lateral communication and interaction between various instruments, enabling various instruments work in conjunction and share information seamlessly in real time. Cloud technology plays a role in enabling real-time data sharing and processing, scientists can access data remotely, collaborate remotely, and utilize computational resources analyze complex datasets. The cartoon scientist is generated by DALLE3.

rapidly test ideas without the need for setting up their own lab, which can dramatically improve efficiency and reduce redundancy, as individuals need not invest in their own laboratories for the purpose of completing a synthesis. This should result in a significant reduction in overall time to discovery, as it removes a significant activation barrier (lab setup).

In contrast, user facilities, especially those run by the government sponsors (e.g., light sources, synchrotrons, and NSRCs), provide capabilities that are not generally available to most individual PIs. Efforts to implement autonomous capabilities are nascent but developing. Still, the fact that skilled instrument scientists are at these facilities enables automation of different aspects of the facility, due to in-house technical knowledge that would be difficult to acquire externally, and can realistically only be shared through these outlets.

Regardless, the development of automated facilities is a prerequisite for the development of autonomous workflows where the whole setup can be optimized. It should further be noted that the optimization, e.g., of synthesis, is not always equitable with a quantitative efficiency gain. For many, if not most experimental workflows, the bottleneck lies not in time taken to perform a single measurement or experiment, but in other points (such as model generation). However, this is not to say, the automated synthesis is not valuable. Rather, it enables rapid iterations to converge on higher-quality solutions that can be found with human expertise alone.[199,200]
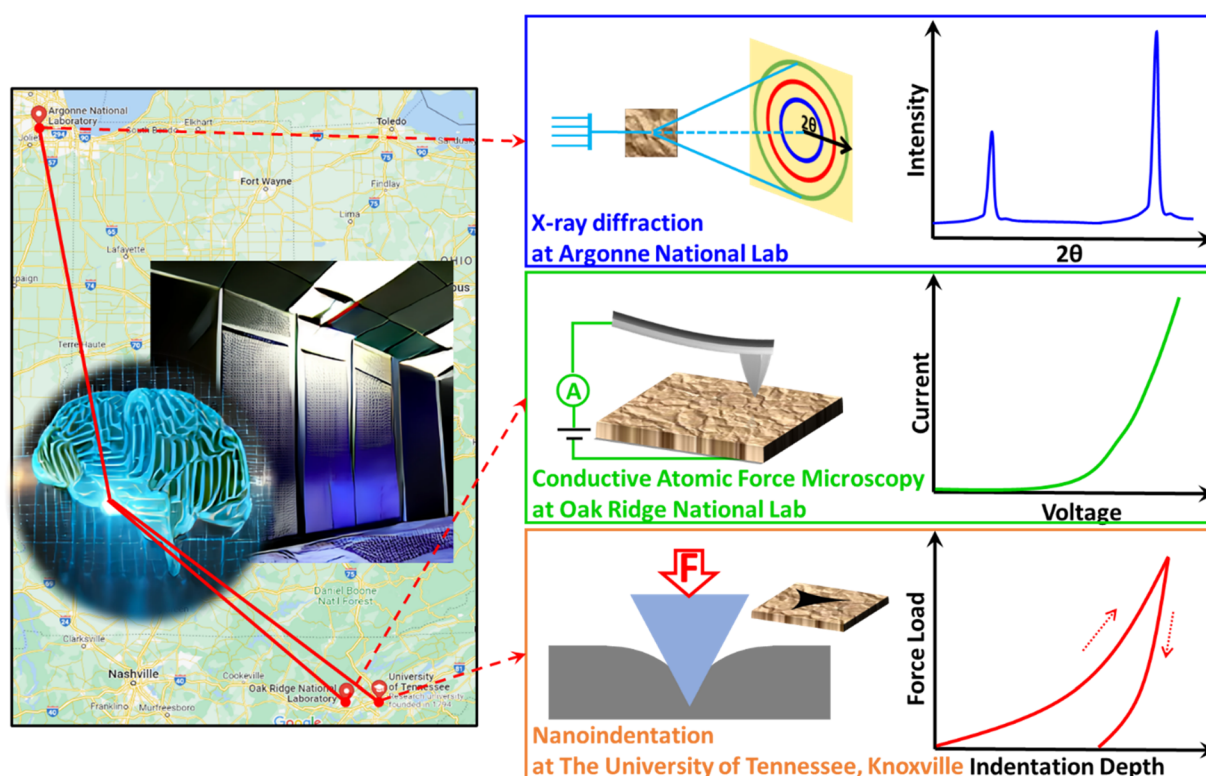
FIG. 8. Graphic representation of the geographically distributed multimodal characterization experiment where the identical combinatorial library is simultaneously imaged via focused x ray at Argonne, nanoindentation at UT Knoxville, and SPM IV at ORNL and the experimental pathways are updated based on results from all three systems.

VII. REQUIREMENTS FOR IMPLEMENTATION OF WORKFLOW

The analysis above allows us to summarize the requirements for broad implementation of scientific workflow design for automated and traditional laboratories. These include the following:

(1) Development of the labs capable of orchestrating pre-defined workflows based on human and robotic agents. These can be purely human-operated lab, purely robotic lab, or human–robot lab where humans perform technical tasks. The associated need is the hyperlanguage that summarizes possible operations and provides access to control parameters. This also includes process monitoring on multiple levels—sample locations, collection of proxy signals during processing, environment monitoring, etc. Here, it is important to have access to full experimental data including both positive and negative outcomes, since purely positive data are often insufficient for optimal ML.

(2) Workflow design based on AI and human decision making, meaning specific series of synthesis and characterization steps based in hyperlanguage. Since physical objects that can be only in one place in one time, workflow will have a directed graph structure (but can form quasi loops when folded on material axis, e.g., for optimization). Note that currently humans both plan workflows and execute them, but these functions can be separated. Key aspect here will be the development of consistent ontologies operating across disparate workflow elements.

(3) Defining domain-specific reward functions that guide workflow development. Why are we running experiments? Is the reward scientific discovery, optimization, or something else (curiosity or empowerment)? Ultimately, we should be able to quantify (in the style of Bellman equation for reinforcement learning)[201] what is the benefit of the specific step in the workflow, and how does it accomplish or affects exploration and exploitation goals.

(4) Integration of reward functions from dissimilar domains, since almost in all cases total reward function will be compounded from multiple intermediate rewards. For example, how does better microscope help us learn physics of specific material? Why would the specific DFT calculation help us understand experimental data?

(5) Creating experimentally falsifiable hypothesis from the domain-specific body of knowledge that can be incorporated in the exploratory part of automated workflows. This is required because workflow design should ideally include a discovery component, and not only optimization. Discovery is effectively extrapolation into unknown domain, and given very high dimensionality, the full space of possible experiments is intractable. Hypotheses provide a way to constrain the space to explore. Note that updating hypotheses based on experimental data is a well-defined problem in Bayesian sense.

(6) Hypothesis generation. In many cases, this is an extrapolation problem and will likely be human-driven (and potentially AI assisted) for the foreseeable future. However, it will be

interesting to explore whether combination of interpolative capabilities of models such as ChatGPT combined with their ability to work with the full volume of text and other data available to mankind will be sufficient for hypothesis generation.

VIII. SUMMARY

Scientific research and discovery are typically organized around workflows, or sequences of the specific actions and experiments targeting specific outcomes. Until now, the workflow design has been highly domain-specific and once established, the workflows remain constant over decades. The disruption of the existing workflows or the introduction of the new ones is typically associated with the emergence of the novel experimental tools. Traditionally, ideation, orchestration, and implementation of the workflows are human-based. The advent of machine learning tools over the last three years has facilitated optimization of human-built workflows, but yet has not led to beyond-human experimentation.

Here, we introduce simple workflows for structural characterization and show that these can be based only on the discovery or weighted by prior knowledge. We suggest several possible strategies for the characterization workflow design. We discuss the increase in complexity for combined imaging-characterization workflows and illustrate the direct and inverse step for design of such workflow.

Finally, we argue that a similar approach can be extended to combined synthesis—imaging-characterization workflows and workflows containing theory in the loop. The workflow design in this case becomes extremely complex and will weigh the latencies, costs, and expected benefits of all steps. We believe that the design of such workflows will require careful analysis of the rewards and the analysis of the value of individual steps. However, the emergence of automated experiments and labs necessitates these developments. Overall, these tools will enable knowledge-based workflow optimization; enable lateral instrumental networks, sequential and parallel orchestration of characterization between dissimilar facilities; and empower distributed research.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Sergei V. Kalinin: Writing – original draft (equal); Writing – review & editing (equal). **Maxim Ziatdinov:** Writing – original draft (lead); Writing – review & editing (lead). **Mahshid Ahmadi:** Writing – original draft (equal); Writing – review & editing (equal). **Ayana Ghosh:** Writing – original draft (equal); Writing – review & editing (equal). **Kevin Roccapriore:** Writing – original draft (equal); Writing – review & editing (equal). **Yongtao Liu:** Writing – original draft (equal); Writing – review & editing (equal). **Rama Vasudevan:** Writing – original draft (lead); Writing – review & editing (lead).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1] S. J. Pennycook, "The impact of STEM aberration correction on materials science," Ultramicroscopy **180**, 22–33 (2017).

[2] C. Gerber and H. P. Lang, "How the doors to the nanoworld were opened," Nat. Nanotechnol. **1**, 3–5 (2006).

[3] P. Sedigh Rahimabadi, M. Khodaei, and K. R. Koswattage, "Review on applications of synchrotron-based X-ray techniques in materials characterization," X-Ray Spectrom. **49**, 348–373 (2020).

[4] G. Martínez-Criado, "Application of micro-and nanobeams for materials science," in Synchrotron Light Sources Free-Electron Lasers, Accelerator Physics, Instrumentation Science Applications (Springer, Cham, 2020), pp. 1719–1753.

[5] C. M. Jeffries et al., "Small-angle x-ray and neutron scattering," Nat. Rev. Methods Primers **1**, 70 (2021).

[6] Y. Tokura, M. Kawasaki, and N. Nagaosa, "Emergent functions of quantum materials," Nat. Phys. **13**, 1056–1068 (2017).

[7] N. Samarth, "Quantum materials discovery from a synthesis perspective," Nat. Mater. **16**, 1068–1076 (2017).

[8] D. N. Basov, R. D. Averitt, and D. Hsieh, "Towards properties on demand in quantum materials," Nat. Mater. **16**, 1077–1088 (2017).

[9] S. V. Kalinin, Y. Kim, D. D. Fong, and A. N. Morozovska, "Surface-screening mechanisms in ferroelectric thin films and their effect on polarization dynamics and domain structures," Rep. Prog. Phys. **81**, 036502 (2018).

[10] Y. Liu et al., "Chemical nature of ferroelastic twin domains in CH3NH3PbI3 perovskite," Nat. Mater. **17**, 1013–1019 (2018).

[11] M. Ahmadi et al., "Exploring anomalous polarization dynamics in organometallic halide perovskites," Adv. Mater. **30**, 1705298 (2018).

[12] F. Dyson, "George green and physics," Phys. World **6**, 33 (1993).

[13] Y. Liu et al., "Ferroic halide perovskite optoelectronics," Adv. Funct. Mater. **31**, 2102793 (2021).

[14] J. N. Wilson, J. M. Frost, S. K. Wallace, and A. Walsh, "Dielectric and ferroic properties of metal halide perovskites," APL Mater. **7**, 010901 (2019).

[15]Y. Liu *et al.*, "Strain–chemical gradient and polarization in metal halide perovskites," Adv. Electron. Mater. **6**, 1901235 (2020).

[16]E. M. Tennyson, T. A. Doherty, and S. D. Stranks, "Heterogeneity at multiple length scales in halide perovskite semiconductors," Nat. Rev. Mater. **4**, 573–587 (2019).

[17]Z. Li *et al.*, "Robot-accelerated perovskite investigation and discovery," Chem. Mater. **32**, 5650–5663 (2020).

[18]J. Piątek *et al.*, "Sustainable Li-ion batteries: Chemistry and recycling," Adv. Energy Mater. **11**, 2003456 (2021).

[19]G. N. Haidemenopoulos, *Physical Metallurgy: Principles and Design* (CRC Press, 2018).

[20]E. P. George, D. Raabe, and R. O. Ritchie, "High-entropy alloys," Nat. Rev. Mater. **4**, 515–534 (2019).

[21]W. Holand and G. H. Beall, *Glass-Ceramic Technology* (John Wiley & Sons, 2019).

[22]M. Dawber, K. M. Rabe, and J. F. Scott, "Physics of thin-film ferroelectric oxides," Rev. Mod. Phys. **77**, 1083–1130 (2005).

[23]J. M. Tarascon and M. Armand, "Issues and challenges facing rechargeable lithium batteries," Nature **414**, 359–367 (2001).

[24]Y. Liu, M. Lorenz, A. V. Ievlev, and O. S. Ovchinnikova, "Secondary ion mass spectrometry (SIMS) for chemical characterization of metal halide perovskites," Adv. Funct. Mater. **30**, 2002201 (2020).

[25]C. Liu, L. Wang, J. Qi, and K. Liu, "Designed growth of large-size 2D single crystals," Adv. Mater. **32**, 2000046 (2020).

[26]S. P. Cramer, *X-Ray Spectroscopy with Synchrotron Radiation* (Springer, 2020).

[27]E. Stach *et al.*, "Autonomous experimentation systems for materials development: A community perspective," Matter **4**, 2702–2726 (2021).

[28]M. Ziatdinov, Y. Liu, K. Kelley, R. Vasudevan, and S. V. Kalinin, "Bayesian active learning for scanning probe microscopy: From Gaussian processes to hypothesis learning," ACS Nano **16**, 13492–13512 (2022).

[29]Y. Jiang *et al.*, "An artificial intelligence enabled chemical synthesis robot exploration optimization nanomaterials," Sci. Adv. **8**, eabo2626 (2022).

[30]D. J. Kowalski, C. M. MacGregor, D.-L. Long, N. L. Bell, and L. Cronin, "Automated library generation and serendipity quantification enables diverse discovery in coordination chemistry," J. Am. Chem. Soc. **145**, 2332–2341 (2023).

[31]D. Caramelli *et al.*, "Discovering new chemistry with an autonomous robotic platform driven by a reactivity-seeking neural network," ACS Cent. Sci. **7**, 1821–1830 (2021).

[32]K. G. Reyes and B. Maruyama, "The machine learning revolution in materials?," MRS Bull. **44**, 530–537 (2019).

[33]D. P. Tabor *et al.*, "Accelerating the discovery of materials for clean energy in the era of smart automation," Nat. Rev. Mater. **3**, 5–20 (2018).

[34]P. Nikolaev *et al.*, "Autonomy in materials research: A case study in carbon nanotube growth," npj Comput. Mater. **2**, 16031 (2016).

[35]B. P. MacLeod *et al.*, "Self-driving laboratory for accelerated discovery of thin-film materials," Sci. Adv. **6**, eaaz8867 (2020).

[36]M. M. Flores-Leonar *et al.*, "Materials acceleration platforms: On the way to autonomous experimentation," Curr. Opin. Green Sustainable Chem. **25**, 100370 (2020).

[37]R. W. Epps, K. C. Felton, C. W. Coley, and M. Abolhasani, "Automated microfluidic platform for systematic studies of colloidal perovskite nanocrystals: Towards continuous nano-manufacturing," Lab Chip **17**, 4040–4047 (2017).

[38]K. Abdel-Latif, F. Bateni, S. Crouse, and M. Abolhasani, "Flow synthesis of metal halide perovskite quantum dots: From rapid parameter space mapping to AI-guided modular manufacturing," Matter **3**, 1053–1086 (2020).

[39]R. W. Epps *et al.*, "Artificial chemist: An autonomous quantum dot synthesis bot," Adv. Mater. **32**, 2001626 (2020).

[40]P. A. Beaucage and T. B. Martin, "The autonomous formulation laboratory: An open liquid handling platform for formulation discovery using x-ray and neutron scattering," Chem. Mater. **35**, 846–852 (2023).

[41]J. M. Gregoire, L. Zhou, and J. A. Haber, "Combinatorial synthesis for AI-driven materials discovery," Nat. Synth. **2**, 493–504 (2023).

[42]H. S. Stein and J. M. Gregoire, "Progress and prospects for accelerating materials science with automated and autonomous workflows," Chem. Sci. **10**, 9640–9649 (2019).

[43]A. G. Kusne *et al.*, "On-the-fly closed-loop materials discovery via Bayesian active learning," Nat. Commun. **11**, 5966 (2020).

[44]F. Rahmanian *et al.*, "One-shot active learning for globally optimal battery electrolyte conductivity," Batteries Supercaps **5**, e202200228 (2022).

[45]N. Choudhary *et al.*, "Autonomous visual detection of defects from battery electrode manufacturing," Adv. Intell. Syst. **4**(12), 2200142 (2022).

[46]Y. Liu *et al.*, "Autonomous scanning probe microscopy with hypothesis learning: Exploring the physics of domain switching in ferroelectric materials," Patterns **4**(3), 100704 (2023).

[47]Y. Liu *et al.*, "Learning the right channel in multimodal imaging: Automated experiment in piezoresponse force microscopy," npj Comput. Mater. **9**, 34 (2023).

[48]A. G. Kusne and A. McDannald, "Scalable multi-agent lab framework for lab optimization," Matter **6**(6), 1880–1893 (2023).

[49]H. S. Stein *et al.*, "From materials discovery to system optimization by integrating combinatorial electrochemistry and data science," Curr. Opin. Electrochem. **35**, 101053 (2022).

[50]Y. Liu *et al.*, "Exploring the relationship of microstructure and conductivity in metal halide perovskites via active learning-driven automated scanning probe microscopy," J. Phys. Chem. Lett. **14**(13), 3352–3359 (2023).

[51]Y. Liu *et al.*, "Automated experiments of local non-linear behavior in ferroelectric materials," Small **18**, 2204130 (2022).

[52]Y. Liu, K. P. Kelley, H. Funakubo, S. V. Kalinin, and M. Ziatdinov, "Exploring physics of ferroelectric domain walls in real time: Deep learning enabled scanning probe microscopy," Adv. Sci. **9**, 2203957 (2022).

[53]A. Benayad *et al.*, "High-throughput experimentation and computational freeway lanes for accelerated battery electrolyte and interface development research," Adv. Energy Mater. **12**, 2102678 (2022).

[54]Y. Liu *et al.*, "Experimental discovery of structure–property relationships in ferroelectric materials via active learning," Nat. Mach. Intell. **4**, 341–350 (2022).

[55]D. Chen *et al.*, "Automating crystal-structure phase mapping by combining deep learning with constraint reasoning," Nat. Mach. Intell. **3**, 812–822 (2021).

[56]Y. Liu *et al.*, "AEcroscoPy: A software-hardware framework empowering microscopy toward automated and autonomous experimentation," arXiv:2312.10281 (2023).

[57]K. Higgins, S. M. Valleti, M. Ziatdinov, S. V. Kalinin, and M. Ahmadi, "Chemical robotics enabled exploration of stability in multicomponent lead halide perovskites via machine learning," ACS Energy Lett. **5**(11), 3426–3436 (2020).

[58]A. Heimbrook, K. Higgins, S. V. Kalinin, and M. Ahmadi, "Exploring the physics of cesium lead halide perovskite quantum dots via Bayesian inference of the photoluminescence spectra in automated experiment," Nanophotonics **10**(8), 1977–1989 (2020).

[59]K. Higgins, M. Ziatdinov, S. V. Kalinin, and M. Ahmadi, "High-throughput study of antisolvents on the stability of multicomponent metal halide perovskites through robotics-based synthesis and machine learning approaches," J. Am. Chem. Soc. **143**(47), 19945–19955 (2021).

[60]S. L. Sanchez *et al.*, "Physics-driven discovery and bandgap engineering of hybrid perovskites," arXiv:2310.06583 (2023).

[61]S. Chen *et al.*, "Exploring the stability of novel wide bandgap perovskites by a robot based high throughput approach," Adv. Energy Mater. **8**(6), 1701543 (2018).

[62]E. Gu *et al.*, "Robot-based high-throughput screening of antisolvents for lead halide perovskites," Joule **4**(8), 1806–1822 (2020).

[63]B. Rohr, H. S. Stein, D. Guevarra, Y. Wang, J. A. Haber, M. Aykol, S. K. Suram, and J. M. Gregoire, "Benchmarking the acceleration of materials discovery by sequential learning," Chem. Sci. **11**, 2696–2706 (2020).

[64]A. Wang, H. Liang, A. McDannald, I. Takeuchi, and A. G. Kusne, "Benchmarking active learning strategies for materials optimization and discovery," Oxford Open Mater. Sci. **2**(1), itac006 (2022).

[65]Q. Liang, A. E. Gongora, Z. Ren, A. Tiihonen, Z. Liu, S. Sun, J. R. Deneault, D. Bash, F. Mekki-Berrada, S. A. Khan, K. Hippalgaonkar, B. Maruyama, K. A. Brown, J. Fisher III, and T. Buonassisi, "Benchmarking the performance of Bayesian optimization across multiple experimental materials science domains," npj Comput. Mater. **7**, 188 (2021).

[66]M. Rief, F. Oesterhelt, B. Heymann, and H. E. Gaub, "Single molecule force spectroscopy on polysaccharides by atomic force microscopy," Science **275**, 1295–1297 (1997).

[67]H. Clausen-Schaumann, M. Seitz, R. Krautbauer, and H. E. Gaub, "Force spectroscopy with single bio-molecules," Curr. Opin. Chem. Biol. **4**, 524–530 (2000).

[68]S. V. Kalinin et al., "Probing the role of single defects on the thermodynamics of electric-field induced phase transitions," Phys. Rev. Lett. **100**, 155703 (2008).

[69]S. Jesse et al., "Direct imaging of the spatial and energy distribution of nucleation centres in ferroelectric materials," Nat. Mater. **7**, 209–215 (2008).

[70]N. Balke et al., "Nanoscale mapping of ion diffusion in a lithium-ion battery cathode," Nat. Nanotechnol. **5**, 749–754 (2010).

[71]S. H. Pan et al., "Imaging the effects of individual zinc impurity atoms on superconductivity in $Bi_2Sr_2CaCu_2O_{8+delta}$," Nature **403**, 746–750 (2000).

[72]P. Roushan et al., "Topological surface states protected from backscattering by chiral spin texture," Nature **460**, 1106–1109 (2009).

[73]R. K. Vasudevan, A. Tselev, A. G. Gianfrancesco, A. P. Baddorf, and S. V. Kalinin, "Atomic-scale electrochemistry on the surface of a manganite by scanning tunneling microscopy," Appl. Phys. Lett. **106**, 143107 (2015).

[74]D. M. Eigler and E. K. Schweizer, "Positioning single atoms with a scanning tunneling microscope," Nature **344**, 524–526 (1990).

[75]R. Danev, H. Yanagisawa, and M. Kikkawa, "Cryo-electron microscopy methodology: Current aspects and future directions," Trends Biochem. Sci. **44**, 837–848 (2019).

[76]B. K. Vainshtein, *Structure Analysis by Electron Diffraction* (Elsevier, 2013).

[77]F. Huo et al., "Beam pen lithography," Nat. Nanotechnol. **5**, 637–640 (2010).

[78]O. Dyck et al., "Atom-by-atom fabrication with electron beams," Nat. Rev. Mater. **4**, 497–507 (2019).

[79]K. M. Roccapriore et al., "Electron-beam induced emergence of mesoscopic ordering in layered $MnPS_3$," ACS Nano **16**, 16713–16723 (2022).

[80]K. Mustonen et al., "Electron-beam manipulation of silicon impurities in single-walled carbon nanotubes," Adv. Funct. Mater. **29**, 1901327 (2019).

[81]S. Maruyama, K. Ouchi, T. Koganezawa, and Y. Matsumoto, "High-throughput and autonomous grazing incidence x-ray diffraction mapping of organic combinatorial thin-film library driven by machine learning," ACS Comb. Sci. **22**, 348–355 (2020).

[82]S. Ament, M. Amsler, D. R. Sutherland, M. C. Chang, D. Guevarra, A. B. Connolly, J. M. Gregoire, M. O. Thompson, C. P. Gomes, and R. B. va. Dover, "Autonomous materials synthesis via hierarchical active learning of nonequilibrium phase diagrams," Sci. Adv. **7**(51), 4930 (2021).

[83]R. Casukhela, S. Vijayan, J. R. Jinschek, and S. R. Niezgoda, "A framework for the optimal selection of high-throughput data collection workflows by autonomous experimentation systems," Integr. Mater. Manuf. Innovation **11**, 557–567 (2022).

[84]K. Hippalgaonkar et al., "Knowledge-integrated machine learning for materials: Lessons from gameplaying and robotics," Nat. Rev. Mater. **8**, 241–260 (2023).

[85]M. Sim, M. G. Vakili, F. Strieth-Kalthoff, H. Hao, R. Hickman, S. Miret, S. Pablo-García, and A. Aspuru-Guzik, "ChemOS 2.0: An orchestration architecture for chemical self-driving laboratories," chemRxiv (2023).

[86]D. Guevarra, K. Kan, Y. Lai, R. J. R. Jones, L. Zhouab, P. Donnelly, M. Richter, H. S. Steinc, and J. M. Gregoire, "Orchestrating nimble experiments across interconnected labs," Digital Discovery **2**, 1806–1812 (2023).

[87]L. Bromig, D. Leiter, A.-V. Mardale, N. von den Eichen, E. Bieringer, and D. Weuster-Botz, "The SiLA 2 Manager for rapid device integration and workflow automation," SoftwareX **17**, 100991 (2022).

[88]R. Tamura, K. Tsuda, and S. Matsuda, "NIMS-OS: An automation software to implement a closed loop between artificial intelligence and robotic experiments in materials science," arXiv:2304.13927 (2023).

[89]A. Musiienko et al., "Deep levels, charge transport and mixed conductivity in organometallic halide perovskites," Energy Environ. Sci. **12**, 1413–1425 (2019).

[90]J. Yang et al., "Understanding the role of cesium on chemical complexity in methylammonium-free metal halide perovskites," Adv. Energy Mater. **13**(33), 2202880 (2022).

[91]A. M. Bran, S. Cox, A. D. White, and P. Schwaller, "ChemCrow: Augmenting large-language models with chemistry tools," arXiv:2304.05376 (2023).

[92]M. C. Ramos, S. S. Michtavy, M. D. Porosoff, and A. D. White, "Bayesian optimization of catalysts with in-context learning," arXiv:2304.05341 (2023).

[93]D. A. Boiko et al., "Autonomous chemical research with large language models," Nature **624**, 7992 (2023).

[94]S. Ament, M. Amsler, D. R. Sutherland, M.-C. Chang, D. Guevarra, A. B. Connolly, J. M. Gregoire, M. O. Thompson, C. P. Gomes, and R. B. van Dover, "Autonomous materials synthesis via hierarchical active learning of nonequilibrium phase diagrams," Sci. Adv. **7**, eabg4930 (2021).

[95]L. Loncan et al., "Hyperspectral pansharpening: A review," IEEE Geosci. Remote Sens. Mag. **3**, 27–46 (2015).

[96]V. Kumar and M. Kumar, "Temporal-spatial-energy resolved advance multidimensional techniques to probe photovoltaic materials from atomistic viewpoint for next-generation energy solutions," Energy Environ. Sci. **14**, 4760–4802 (2021).

[97]J. Hidalgo, A. F. Castro-Méndez, and J. P. Correa-Baena, "Imaging and mapping characterization tools for perovskite solar cells," Adv. Energy Mater. **9**, 1900444 (2019).

[98]J. Leszczynski, *Practical Aspects of Computational Chemistry* (Springer, 2012).

[99]S. Yip, *Handbook of Materials Modeling* (Springer-Verlag, 2005).

[100]V. E. Panin, *Foundations of Physical Mesomechanics* (Springer, 1998), Vol. 1, pp. 5–20.

[101]G. Lu, *Handbook of Theoretical and Computational Nanotechnology* (American Scientific Publishers, 2004).

[102]K. Matouš, M. G. Geers, V. G. Kouznetsova, and A. Gillman, "A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials," J. Comput. Phys. **330**, 192–220 (2017).

[103]B. Chopard, J. L. Falcone, P. Kunzli, L. Veen, and A. Hoekstra, "Multiscale modeling: Recent progress and open questions," Multiscale Multidiscip. Model., Exp. Des. **1**, 57–68 (2018).

[104]J. Fish, G. J. Wagner, and S. Keten, "Mesoscopic and multiscale modelling in materials," Nat. Mater. **20**, 774–786 (2021).

[105]D. Bishara, Y. Xie, W. K. Liu, and S. Li, "A state-of-the-art review on machine learning-based multiscale modeling, simulation, homogenization and design of materials," Arch. Comput. Methods Eng. **30**, 191–222 (2023).

[106]G. Lu et al., "From electrons to finite elements: A concurrent multiscale approach for metals," Phys. Rev. B **73**, 024108 (2006).

[107]D. L. McDowell and G. B. Olson, "Concurrent design of hierarchical materials and structures," Sci. Model. Simul. **15**, 207–240 (2008).

[108]E. B. Tadmor et al., "Hierarchical modeling in the mechanics of materials," Int. J. Solids Struct. **37**(1–2), 379–389 (2000).

[109]A. Ramasubramaniam and E. A. Carter, "Coupled quantum–atomistic and quantum–continuum mechanics methods in materials research," MRS Bull. **32**, 913–918 (2007).

[110]N. Choly et al., "Multiscale simulations in simple metals: A density-functional-based methodology," Phys. Rev. B **71**, 094101 (2005).

[111]G. Csanyi et al., ""Learn on the Fly": A hybrid classical and quantum-mechanical molecular dynamics simulation," Phys. Rev. Lett. **93**, 175503 (2004).

[112]J. Behler, "Perspective: Machine learning potentials for atomistic simulations," J. Chem. Phys. **145**, 170901 (2016).

[113]C. W. Rosenbrock, K. Gubaev, A. V. Shapeev, L. B. Pártay, N. Bernstein, G. Csányi, and G. L. W. Hart, "Machine-learned interatomic potentials for alloys and alloy phase diagrams," npj Comput. Mater. **7**, 24 (2021).

[114]G. Sivaraman, A. N. Krishnamoorthy, M. Baur, C. Holm, M. Stan, G. Csányi, C. Benmore, and Á. Vázquez-Mayagoitia, "Machine-learned interatomic potentials by active learning: Amorphous and liquid hafnium dioxide," npj Comput. Mater. **6**, 104 (2020).

[115]B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: A review of Bayesian optimization," Proc. IEEE **104**, 148–175 (2015).

[116]T. Mueller, A. Hernandez, and C. Wang, "Machine learning for interatomic potential models," J. Chem. Phys. **152**, 050902 (2020).

[117]G. P. Wellawatte, A. Seshadri, and A. D. White, "Model agnostic generation of counterfactual explanations for molecules," chemRxiv (2021).

[118]V. Botu, R. Batra, J. Chapman, and R. Ramprasad, "Machine learning force fields: Construction, validation, and outlook," J. Phys. Chem. C **121**, 511–522 (2017).

15 April 2024 16:13:56

[119] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, "Parameter space compression underlies emergent theories and predictive models," Science **342**, 604–607 (2013).

[120] J. Carrasquilla and R. G. Melko, "Machine learning phases of matter," Nat. Phys. **13**, 431–434 (2017).

[121] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," Found. Trends® Mach. Learn. **12**, 307–392 (2019).

[122] S. V. Kalinin, J. J. Steffes, Y. T. Liu, B. D. Huey, and M. Ziatdinov, "Disentangling ferroelectric domain wall geometries and pathways in dynamic piezoresponse force microscopy via unsupervised machine learning," Nanotechnology **33**, 055707 (2022).

[123] R. Ignatans et al., "Latent mechanisms of polarization switching from in situ electron microscopy observations," Adv. Funct. Mater. **32**, 2100271 (2022).

[124] S. V. Kalinin et al., "Deep Bayesian local crystallography," npj Comput. Mater. **7**, 181 (2021).

[125] P. N. Loxley and K.-W. Cheung, "A dynamic programming algorithm for finding an optimal sequence of informative measurements," Entropy **25**, 251 (2023).

[126] M. A. Ziatdinov et al., "Hypothesis learning in automated experiment: Application to combinatorial materials libraries," Adv. Mater. **34**(20), 2201345 (2022).

[127] D. Silver et al., "Mastering the game of go without human knowledge," Nature **550**, 354–359 (2017).

[128] K. Shao, Z. Tang, Y. Zhu, N. Li, and D. Zhao, "A survey of deep reinforcement learning in video games," arXiv:1912.10944 (2019).

[129] C. Berner et al., "Dota 2 with large scale deep reinforcement learning," arXiv:1912.06680 (2019).

[130] J. Skalse, N. Howe, D. Krasheninnikov, and D. Krueger, "Defining and characterizing reward gaming," Adv. Neural Inf. Process. Syst. **35**, 9460–9471 (2022); available at https://proceedings.neurips.cc/paper_files/paper/2022/hash/3d719fee332caa23d5038b8a90e81796-Abstract-Conference.html

[131] C. Laidlaw, S. Singhal, and A. Dragan, "Preventing reward hacking with occupancy measure regularization," in The Second Workshop on New Frontiers in Adversarial Machine Learning, International Conference on Machine Learning (2023).

[132] Y. Liu, K. P. Kelley, H. Funakubo, S. V. Kalinin, and M. Ziatdinov, "Exploring physics of ferroelectric domain walls in real time: Deep learning enabled scanning probe microscopy," Adv. Sci. **9**(31), 2203957 (2022).

[133] Y. Liu et al., "Disentangling electronic transport and hysteresis at individual grain boundaries in hybrid perovskites via automated scanning probe microscopy," ACS Nano **17**(10), 9647–9657 (2023).

[134] Y. Liu, M. A. Ziatdinov, R. K. Vasudevan, and S. V. Kalinin, "Explainability and human intervention in autonomous scanning probe microscopy," Patterns **4**(11), 100858 (2023).

[135] J. Morbach, A. Yang, and W. Marquardt, "OntoCAPE—A large-scale ontology for chemical process engineering," Eng. Appl. Artif. Intell. **20**, 147–161 (2007).

[136] K. S. Aggour et al., "Compound knowledge graph-enabled AI assistant for accelerated materials discovery," Integr. Mater. Manuf. Innovation **11**, 467–478 (2022).

[137] M. J. Statt, B. A. Rohr, D. Guevarra, S. K. Suram, and J. M. Gregoire, "The materials experiment knowledge graph," Digital Discovery **2**, 909–914 (2023).

[138] S. V. Kalinin, A. Ghosh, R. Vasudevan, and M. Ziatdinov, "From atomically resolved imaging to generative and causal models," Nat. Phys. **18**, 1152–1160 (2021).

[139] R. Potyrailo, K. Rajan, K. Stoewe, I. Takeuchi, B. Chisholm, and H. Lam, "Combinatorial and high-throughput screening of materials libraries: Review of state of the art," ACS Comb. Sci. **13**(6), 579–633 (2011).

[140] W. F. Maier, K. Stowe, and S. Sieg, " Combinatorial and high-throughput materials science," Angew. Chem. **46**(32), 6016–6067 (2007).

[141] W. F. Maier, "Early years of high-throughput experimentation and combinatorial approaches in catalysis and materials science," ACS Comb. Sci. **21**, 437–444 (2019).

[142] M. L. Green, C. L. Choi, J. R. Hattrick-Simpers, A. M. Joshi, I. Takeuchi, S. C. Barron, E. Campo, T. Chiang, S. Empedocles, J. M. Gregoire, A. G. Kusne, J. Martin, A. Mehta, K. Persson, Z. Trautt, J. Van Duren, and A. Zakutayev, "Fulfilling the promise of the materials genome initiative with high-throughput experimental methodologies," Appl. Phys. Rev. **4**, 011105 (2017).

[143] J. N. Cawse, "Experimental strategies for combinatorial and high-throughput materials development," Acc. Chem. Res. **34**, 213–221 (2001).

[144] A. Ludwig, "Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods," npj Comput. Mater. **5**, 70 (2019).

[145] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," Nat. Mater. **12**, 191–201 (2013).

[146] J.-P. Correa-Baena, K. Hippalgaonkar, J. van Duren, S. Jaffer, V. R. Chandrasekhar, V. Stevanovic, C. Wadia, S. Guha, and T. Buonassisi, "Accelerating materials development via automation, machine learning, and high-performance computing," Joule **2**, 1410–1420 (2018).

[147] W. Y. Wang, J. Li, W. Liu, and Z.-K. Liu, "Integrated computational materials engineering for advanced materials: A brief review," Comput. Mater. Sci. **158**, 42–48 (2019).

[148] P. Xu, X. Ji, M. Li, and W. Lu, "Small data machine learning in materials science," npj Comput. Mater. **9**, 42 (2023).

[149] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "The Materials Project: A materials genome approach to accelerating materials innovation," APL Mater. **1**, 011002 (2013).

[150] D. Hicks, M. J. Mehl, E. Gossett, C. Toher, O. Levy, R. M. Hanson, G. Hart, and S. Curtarolo, "The AFLOW library of crystallographic prototypes: Part 2," Comput. Mater. Sci. **161**, S1–S1011 (2019).

[151] J. E. Saal, S. K, M. Aykol, B. Meredig, and C. Wolverton, "Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD)," JOM **65**, 1501–1509 (2013).

[152] M. J. Mehl, D. Hicks, C. Toher, O. Levy, R. M. Hanson, G. L. W. Hart, and S. Curtarolo, "The AFLOW library of crystallographic prototypes: Part 1," Comput. Mater. Sci. **136**, S1–S828 (2017).

[153] S. Kirklin, J. E. Saal, B. Meredig, B. A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, "The open quantum materials database (OQMD): Assessing the accuracy of DFT formation energies," npj Comput. Mater. **1**, 15010 (2015).

[154] C. Draxl and M. Scheffler, "The NOMAD laboratory: From data sharing to artificial intelligence," J. Phys. Mater. **2**, 036001 (2019).

[155] A. Jain, J. Montoya, S. Dwaraknath, N. E. R. Zimmermann, J. Dagdelen, M. Horton, P. Huck, D. Winston, S. Cholia, S. P. Ong, and K. Persson, The Materials Project: Accelerating Materials Design through Theory-Driven Data and Tools (Springer, 2020).

[156] A. Ghosh, F. Ronning, S. M. Nakhmanson, and J.-X. Zhu, "Machine learning study of magnetism in uranium-based compounds," Phys. Rev. Mater. **4**, 064414 (2020).

[157] A. Ghosh, L. Louis, K. K. Arora, B. C. Hancock, J. F. Krzyzaniak, P. Meenan, S. Nakhmanson, and G. P. F. Wood, "Assessment of machine learning approaches for predicting the crystallization propensity of active pharmaceutical ingredients," CrystEngComm **21**, 1215–1223 (2019).

[158] R. Batra, "Accurate machine learning in materials science facilitated by using diverse data sources," Nature **589**(7843), 524–525 (2021).

[159] C. Toher, C. Oses, D. Hicks, E. Gossett, F. Rose, P. Nath, D. Usanmaz, D. C. Ford, E. Perim, C. E. Calderon, J. J. Plata, Y. Lederer, M. Jahnátek, W. Setyawan, S. Wang, J. Xue, K. Rasch, R. V. Chepulskii, R. H. Taylor, G. Gomez, H. Shi, A. R. Supka, R. Al Rahal Al Orabi, P. Gopal, F. T. Cerasoli, L. Liyanage, H. Wang, I. Siloi, L. A. Agapito, C. Nyshadham, G. L. W. Hart, J. Carrete, F. Legrain, N. Mingo, E. Zurek, O. Isayev, A. Tropsha, S. Sanvito, R. M. Hanson, I. Takeuchi, M. J. Mehl, A. N. Kolmogorov, K. Yang, P. D'Amico, A. Calzolari, M. Costa, R. De Gennaro, M. B. Nardelli, M. Fornari, O. Levy, and S. Curtarolo, The AFLOW Fleet for Materials Discovery (Springer International Publishing, 2020), pp. 1785–1812.

[160] G. Pilania, A. Ghosh, S. T. Hartman, R. Mishra, C. R. Stanek, and B. P. Uberuaga, "Anion order in oxysulfide perovskites: Origins and implications," npj Comput. Mater. **6**, 71 (2020).

[161] K. M. Merz, Jr., R. Amaro, Z. Cournia, M. Rarey, T. Soares, A. Tropsha, H. A. Wahab, and R. Wang, "Method and data sharing and reproducibility of scientific results," J. Chem. Inf. Model. **60**, 5868–5869 (2020).

[162] L. Chen, G. Pilania, R. Batra, T. D. Huan, C. Kim, C. Kuenneth, R. Ramprasad, L. Chen et al., "Polymer informatics: Current status and critical next steps," Mater. Sci. Eng., R **144**, 100595 (2021).

15 April 2024 16:13:56

163B. Meredig *et al.*, "Combinatorial screening for new materials in unconstrained composition space with machine learning," Phys. Rev. B **89**, 094104 (2014).

164N. Artrith, K. T. Butler, F.-X. Coudert, S. Han, O. Isayev, A. Jain, and A. Walsh, "Best practices in machine learning for chemistry," Nat. Chem. **13**, 505–508 (2021).

165O. Isayev, A. Tropsha, and S. Curtarolo, *Materials Informatics: Methods, Tools, and Applications* (John Wiley & Sons, 2019).

166O. Isayev, M. Popova, and A. Tropsha, "Methods, systems and non-transitory computer readable media for automated design of molecules with desired properties using artificial intelligence," U.S. patent WO2019018780A1 (2020).

167R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine learning in materials informatics: Recent applications and prospects," npj Comput. Mater. **3**, 54 (2017).

168J. Schmidt, M. R. Marques, S. Botti, and M. A. Marques, "Recent advances and applications of machine learning in solid-state materials science," npj Comput. Mater. **5**, 83 (2019).

169J. C. Snyder, M. Rupp, K. Hansen, K.-R. Muller, and K. Burke, "Finding density functionals with machine learning," Phys. Rev. Lett. **108**, 253002 (2012).

170V. Stanev *et al.*, "Machine learning modeling of superconducting critical temperature," npj Comput. Mater. **4**, 29 (2018).

171S. Sanvito, M. Žic, J. Nelson, T. Archer, C. Oses, and S. Curtarolo, *Machine Learning and High-Throughput Approaches to Magnetism* (Springer International Publishing, 2020), pp. 351–373.

172A. Ghosh, M. Ziatdinov, O. Dyck, B. Sumpter, and S. V. Kalinin, "Bridging microscopy with molecular dynamics and quantum simulations: An atom AI based pipeline," npj Comput. Mater. **8**, 74 (2022).

173E. J. Lingerfelt, A. Belianinov, E. Endeve, O. Ovchinnikov, S. Somnath, J. M. Borreguero, N. Grodowitz, B. Park, R. K. Archibald, C. T. Symons, S. V. Kalinin, O. E. B. Messer, M. Shankar, and S. Jesse, "BEAM: A computational workflow system for managing and modeling material characterization data in HPC environments," Procedia Comput. Sci. **80**, 2276–2280 (2016).

174E. Schwenker *et al.*, "Ingrained: An automated framework for fusing atomic-scale image simulations into experiments," Small **18**(19), 2102960 (2022).

175J. Madsen and T. Susi, "The abTEM code: Transmission electron microscopy from first principles," Open Res. Europe **1**, 24 (2021).

176S. V. Kalinin, M. Ziatdinov, J. Hinkle, S. Jesse, A. Ghosh, K. P. Kelley, A. R. Lupini, B. G. Sumpter, and R. K. Vasudevan, "Automated and autonomous experiments in electron and scanning probe microscopy," ACS Nano **15**, 12604–12627 (2021).

177O. Dyck, S. Kim, E. Jimenez-Izal, A. N. Alexandrova, S. V. Kalinin, and S. Jesse, "Building structures atom by atom via electron beam manipulation," Small **14**, 1801771 (2018).

178O. Dyck, M. Yoon, L. Zhang, A. R. Lupini, J. L. Swett, and S. Jesse, "Doping of Cr in graphene using electron beam manipulation for functional defect engineering," ACS Appl. Nano Mater. **3**, 10855–10863 (2020).

179O. Dyck, L. Zhang, M. Yoon, J. L. Swett, D. Hensley, C. Zhang, P. D. Rack, J. D. Fowlkes, A. R. Lupini, and S. Jesse, "Doping transition-metal atoms in graphene for atomic-scale tailoring of electronic, magnetic, and quantum topological properties," Carbon **173**, 205–214 (2021).

180O. Dyck, C. Zhang, P. D. Rack, J. D. Fowlkes, B. Sumpter, A. R. Lupini, S. V. Kalinin, and S. Jesse, "Electron-beam introduction of heteroatomic Pt–Si structures in graphene," Carbon **161**, 750–757 (2020).

181J. Lin, Y. Zhang, W. Zhou, and S. T. Pantelides, "Structural flexibility and alloying in ultrathin transition-metal chalcogenide nanowires," ACS Nano **10**, 2782–2790 (2016).

182X. Sang, X. Li, W. Zhao, J. Dong, C. M. Rouleau, D. B. Geohegan, F. Ding, K. Xiao, and R. R. Unocic, "In situ edge engineering in two-dimensional transition metal dichalcogenides," Nat. Commun. **9**, 2051 (2018).

183T. Lehnert, M. Ghorbani-Asl, J. Köster, Z. Lee, A. V. Krasheninnikov, and U. Kaise, "Electron-beam-driven structure evolution of single-layer MoTe$_2$ for quantum devices," ACS Appl. Nano Mater. **2**, 3262–3270 (2019).

184W. Zhou, X. Zou, S. Najmaei, Z. Liu, Y. Shi, J. Kong, J. Lou, P. M. Ajayan, B. I. Yakobson, and J.-C. Idrobo, "Intrinsic structural defects in monolayer molybdenum disulfide," Nano Lett. **13**, 2615–2622 (2013).

185J. Hong, Z. Hu, M. Probert, K. Li, D. Lv, X. Yang, L. Gu, N. Mao, Q. Feng, L. Xie, and J. Zhang, "Exploring atomic defects in molybdenum disulfide monolayers," Nat. Commun. **6**, 6293 (2015).

186T. Susi, J. Kotakoski, D. Kepaptsoglou, C. Mangler, T. C. Lovejoy, O. L. Krivanek, R. Zan, U. Bangert, P. Ayala, J. C. Meyer, and Q. Ramasse, "Silicon–carbon bond inversions driven by 60-keV electrons in graphene," Phys. Rev. Lett. **113**, 115501 (2014).

187Z. Yang, D. Lin. Y, J. Lee, W. Ren, H.-M. Cheng, H. Ye, S. T. Pantelides, S. J. Pennycook, and M. F. Chisholm, "Direct observation of atomic dynamics and silicon doping at a topological defect in graphene," Angew. Chem. **126**, 9054–9058 (2014).

188J. Lee, W. Zhou, S. J. Pennycook, J.-C. Idrobo, and S. T. Pantelides, "Direct visualization of reversible dynamics in a Si6 cluster embedded in a graphene pore," Nat. Commun. **4**, 1650 (2013).

189A. W. Robertson, G.-D. Lee, K. He, E. Yoon, A. I. Kirkland, and J. H. Warner, "The role of the bridging atom in stabilizing odd numbered graphene vacancies," Nano Lett. **14**, 3972–3980 (2014).

190A. W. Robertson, G.-D. Lee, K. He, E. Yoon, A. I. Kirkland, and J. H. Warner, "Stability and dynamics of the tetravacancy in graphene," Nano Lett. **14**, 1634–1642 (2014).

191Z. He, K. He, A. W. Robertson, A. I. Kirkland, D. Kim, J. Ihm, E. Yoon, G.-D. Lee, and J. H. Warner, "Atomic structure and dynamics of metal dopant pairs in graphene," Nano Lett. **14**, 3766–3772 (2014).

192Y. Liu, M. Ziatdinov, and S. V. Kalinin, "Exploring causal physical mechanisms via non-gaussian linear models and deep kernel learning: Applications for ferroelectric domain structures," ACS Nano **16**, 1250–1259 (2022).

193A. Ghosh, G. Palanichamy, D. P. Trujillo, M. Shaikh, and S. Ghosh, "Insights into cation ordering of double perovskite oxides from machine learning and causal relations," Chem. Mater. **34**, 7563–7578 (2022).

194M. Ziatdinov, C. T. Nelson, X. Zhang, R. K. Vasudevan, E. Eliseev, A. N. Morozovska, I. Takeuchi, and S. V. Kalinin, "Causal analysis of competing atomistic mechanisms in ferroelectric materials from high-resolution scanning transmission electron microscopy data," npj Comput. Mater. **6**, 127 (2020).

195B. N. Slautin, U. Pratiush, I. N. Ivanov, Y. Liu, R. Pant, X. Zhang, I. Takeuchi, M. A. Ziatdinov, and S. V. Kalinin, "Multimodal co-orchestration for exploring structure-property relationships in combinatorial libraries via multi-task Bayesian optimization." arXiv preprint arXiv:2402.02198 (2024).

196J. D. Martin, *Solid State Insurrection: How the Science of Substance Made American Physics Matter* (University of Pittsburgh Press, 2018).

197C. C. Mody, *The Long Arm of Moore's Law: Microelectronics and American Science* (MIT Press, 2016).

198F. Rahmanian *et al.*, "Enabling modular autonomous feedback-loops in materials science through hierarchical experimental laboratory automation and orchestration," Adv. Mater. Interfaces **9**, 2101987 (2022).

199K. J. Kanarik, W. T. Osowiecki, Y. Lu, D. Talukder, N. Roschewsky, S. N. Park, M. Kamon, D. M. Fried, and R. A. Gottscho, "Human–machine collaboration for improving semiconductor process development," Nature **616**, 707–711 (2023).

200K. L. Snapp, B. Verdier, A. Gongora, S. Silverman, A. D. Adesiji, E. F. Morgan, T. J. Lawton, E. Whiting, and K. A. Brown, "Autonomous discovery of tough structures," arXiv:2308.02315 (2023).

201W. B. Powell, *Reinforcement Learning and Stochastic Optimization: A unified framework for sequential decisions* (John Wiley & Sons, 2022).

15 April 2024 16:13:56