ACS NANO

www.acsnano.org

REVIEW

Reducing Time to Discovery: Materials and Molecular Modeling, Imaging, Informatics, and Integration

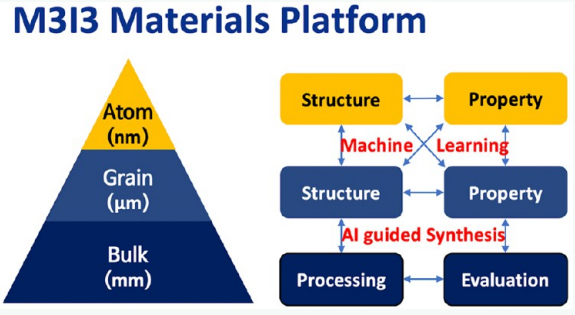
Seungbum Hong,* Chi Hao Liow, Jong Min Yuk, Hye Ryung Byon, Yongsoo Yang, EunAe Cho, Jiwon Yeom, Gun Park, Hyeonmuk Kang, Seunggu Kim, Yoonsu Shim, Moony Na, Chaehwa Jeong, Gyuseong Hwang, Hongjun Kim, Hoon Kim, Seongmun Eom, Seongwoo Cho, Hosun Jun, Yongju Lee, Arthur Baucour, Kihoon Bang, Myungjoon Kim, Seokjung Yun, Jeongjae Ryu, Youngjoon Han, Albina Jetybayeva, Pyuck-Pa Choi, Joshua C. Agar, Sergei V. Kalinin, Peter W. Voorhees, Peter Littlewood, and Hyuck Mo Lee

Cite This: *ACS Nano* 2021, 15, 3971−3995

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Multiscale and multimodal imaging of material structures and properties provides solid ground on which materials theory and design can flourish. Recently, KAIST announced 10 flagship research fields, which include KAIST Materials Revolution: Materials and Molecular Modeling, Imaging, Informatics and Integration (M3I3). The M3I3 initiative aims to reduce the time for the discovery, design and development of materials based on elucidating multiscale processing−structure−property relationship and materials hierarchy, which are to be quantified and understood through a combination of machine learning and scientific insights. In this review, we begin by introducing recent progress on related initiatives around the globe, such as the Materials Genome Initiative (U.S.), Materials Informatics (U.S.), the Materials Project (U.S.), the Open Quantum Materials Database (U.S.), Materials Research by Information Integration Initiative (Japan), Novel Materials Discovery (E.U.), the NOMAD repository (E.U.), Materials Scientific Data Sharing Network (China), Vom Materials Zur Innovation (Germany), and Creative Materials Discovery (Korea), and discuss the role of multiscale materials and molecular imaging combined with machine learning in realizing the vision of M3I3. Specifically, microscopies using photons, electrons, and physical probes will be revisited with a focus on the multiscale structural hierarchy, as well as structure−property relationships. Additionally, data mining from the literature combined with machine learning will be shown to be more efficient in finding the future direction of materials structures with improved properties than the classical approach. Examples of materials for applications in energy and information will be reviewed and discussed. A case study on the development of a Ni−Co−Mn cathode materials illustrates M3I3's approach to creating libraries of multiscale structure−property−processing relationships. We end with a future outlook toward recent developments in the field of M3I3.



KEYWORDS: *M3I3, materials and molecular modeling, materials imaging, materials informatics, machine learning, materials integration, Li-ion battery, KAIST*

The scientific method, a system by which observation, experimentation, and reason are employed in forming and testing hypotheses and theories, began to advance near the end of the Renaissance period. Descartes promoted science by first questioning everything and then constructing a theory based upon sound observational evidence. Materials science is no exception in the sense that the observation of structural features, order parameters fields, and essential

Received: January 8, 2021
Accepted: February 1, 2021
Published: February 12, 2021

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

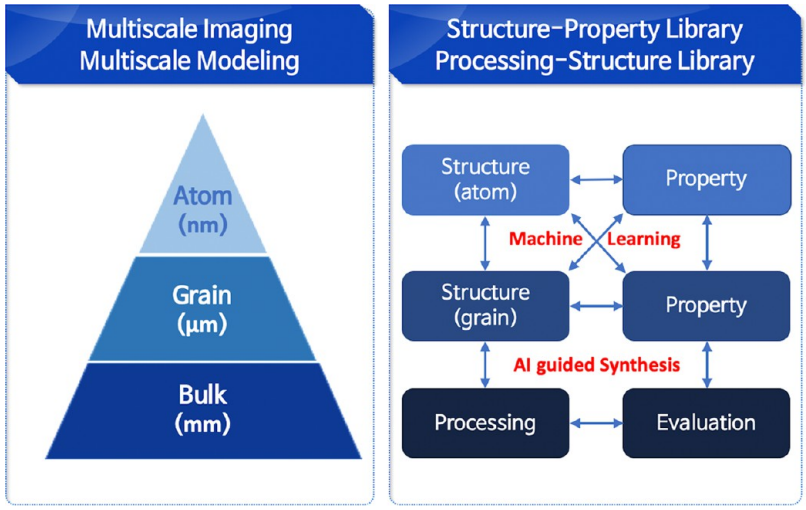


Figure 1. Schematic diagram of M3I3 Flagship Project. This project aims to achieve the seamless integration of the multiscale "structure–property" and "processing–property" relationships *via* materials modeling, imaging, and machine learning. With the capability of artificial intelligence (AI)-guided automatic synthesis, M3I3 will provide expedited development of new materials in the near future.

materials properties provides solid ground on which materials theory and design can flourish.

Materials science and engineering have evolved significantly from empirical trial-and-error approaches, searching and optimization routines, and treasure hunting modes, which are characterized as serendipitous discovery-led development using processing recipes coupled with available technologies. With advances in computational materials science and nanotechnology, "Materials by Design" has emerged as the central paradigm of materials science and engineering in the 2000s.[1]

What is meant by Materials by Design is a synthetic systems view that integrates accumulated knowledge derived from a reductionist analysis while replacing discovery-based R&D with a far more effective and efficient design-based approach. Materials engineers define a quantitative set of properties as objectives, which together results in a desired material performance within the economic boundary conditions. Next, they conduct a systems analysis to identify and prioritize key structure–property and processing–structure relationships. Central to the materials design approach is a powerful logical structure connecting the four elements of processing, structure, property, and performance. The deductive, cause and effect logic of reductionist materials science flows from processing to performance, whereas the inductive logic of systems engineering flows in the reverse direction from performance to processing, thereby enabling materials engineers to design a specific processing recipe to yield materials with the desired sets of properties and performance.[2] However, limited time and cost constraints of the full materials development cycle in industry and the lack of complete information on structure–property as well as processing–property relationships pose great challenges for achieving the vision of Materials by Design.

To tackle the challenges described above, the "Materials Genome Initiative (MGI)" was initiated in 2011 in the U.S. with the idea of using chemical elements, phases, and processes as the genome for designing, manufacturing, and deploying inexpensive materials and materials-based technologies significantly faster than before.[3] Since the birth of MGI, many related initiatives have been launched around the globe, such as the Materials Research by Information Integration Initiative in Japan,[4] Novel Materials Discovery in the E.U.,[5] Vom Materials Zur Innovation in Germany,[6] Materials Scientific Data Sharing Network in China,[7] and Creative Materials Discovery in Korea.[8]

Exploring the large number of materials and predicting their desired properties have enabled the creation of searchable databases for rapidly selecting candidates to be used in experimental studies.[9,10] However, this theory-driven approach has been met with a few important challenges. First, a closer tie-in with experiment is needed to improve the accuracy of predictions. Second, many interesting material properties are defined over different length scales and not just the atomic scale; thus, the computational cost of first-principles methods becomes prohibitive.[11] This in turn necessitates introduction of the mesoscopic averaged models, which requires a large number of phenomenological parameters.

In 2018, KAIST published the "Future Report of KAIST," where 10 flagship research projects were selected through rigorous screening work by all of the faculty members at KAIST.[12] Among the 10 projects, Materials and Molecular Modeling, Imaging, Informatics and Integration (M3I3) was selected as the only one in the field of materials science and engineering (see Figure 1). In 2019, KAIST launched two Global Singularity Projects where M3I3 was selected as the top contender between the two.[13,14]

In this review, we briefly revisit the history of materials science and engineering to understand how materials were discovered and developed, followed by an overview of the role of materials imaging in materials discovery and development. We then cover the emergence of high-throughput screening using density functional theory and machine learning within the framework of Materials by Design. Recent advancements in multimodal and multiscale materials imaging at user facilities around the globe are discussed, which inspired the birth of the M3I3 initiative. Furthermore, we present an example of M3I3 application to rechargeable battery materials. Finally, the role of multiscale materials and molecular imaging combined with machine learning for realizing the vision of M3I3 is presented. We end with a future outlook toward developments and the major challenges of M3I3.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971–3995

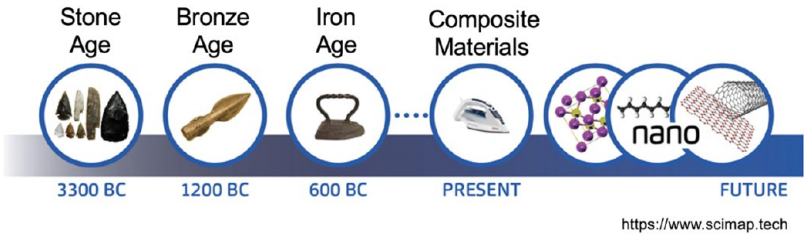


Figure 2. Evolution of historical epochs named by the materials of each era. Adapted with permission from ref 16. Copyright 2018 SciMap.

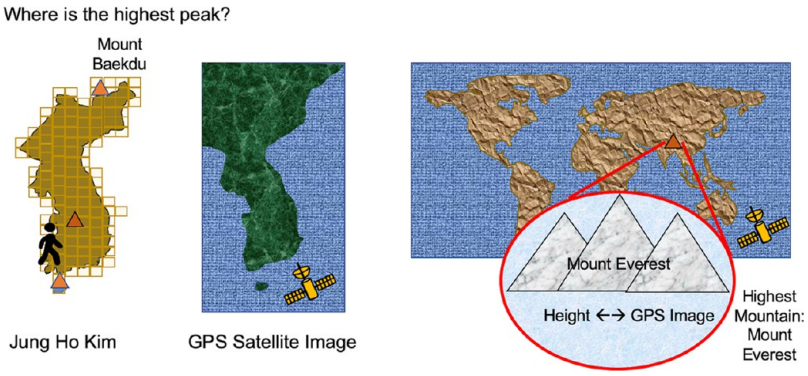


Figure 3. Schematic diagram showing Jung Ho Kim traveling around the country to draw Daedongyeo Map, and a GPS satellite images of the Korean Peninsula and the world. The inset is a schematic image of Mount Everest.

HISTORY OF MATERIALS SCIENCE AND ENGINEERING

How do we divide our history into important segments? By the material of choice. The most advanced material of a given era is often a defining point in history, which is apparent in the naming of historical epochs such as the Stone Age, Bronze and Iron Ages, and the Silicon Age.[15,16] Currently, we can say that we are living in the Composite Nanomaterials Age (see Figure 2). Since we are now entering into the fourth industrial revolution, where all devices are connected and embedded in buildings, vehicles, and even human bodies, more diverse and custom-tailored materials will be needed beyond those that are currently available with mass production-compatible materials.

Then the following question arises: How can we discover and develop materials in response to the challenges of the fourth industrial revolution, the increase in national protectionism due to the COVID-19 era, and the rise of artificial intelligence, open data policies, and user facilities throughout the world? Before thinking about the answer to this question, we will revisit the history of materials science and engineering. Modern materials science evolved directly from metallurgy, ceramics, and solid-state physics/chemistry. Gibbs made a major breakthrough in the understanding of materials when he found that the thermodynamic properties of various phases are related to the physical properties of a material.[17]

Modern materials science had emerged by the second industrial revolution, which was led by the mass production of steam and internal combustion engines, the Second World War, Moore's law, and the Space Race as well as the third industrial revolution led by the Internet and communication technologies. Materials science has driven, and been driven by, the development of revolutionary technologies including but not limited to superalloys, rubbers, plastics, semiconductors, optical fibers, functional ceramics, high-entropy alloys, and biomaterials to name a few.[18,19]

MATERIALS IMAGING: STRUCTURE *VS* PROPERTY

To expedite the development of materials, materials scientists and engineers have worked together to understand the structure and property relationships that underlie the economic value of materials. For example, Sorby revealed internal microstructures by etching metal samples with acid and observed these samples with an optical microscope to correlate the structures of these materials with their properties and performance.[20] This technique enabled scientists to establish well-known structure−property relationships such as the Hall−Petch equation, which relates grain size with the yield strength of mild steel.[21]

It is notable that the Nobel Prize in Physics in 1986 was one-half awarded to Ernst Ruska for his fundamental work in electron optics and for the design of the electron microscope; the other half, jointly to Gerd Binnig and Heinrich Rohrer for their design of the scanning tunneling microscope (STM) and the atomic force microscope (AFM).[22] While transmission electron microscopy (TEM) was invented in the 1930s,[23] it only became widely available many decades later due to the challenges of economics, scale, and manufacturing of these precision technologies. The big driving forces for commercialization of electron microscopy were the semiconductor industries that need to image their integrated circuit (IC)/CMOS devices and hard disk drives (HDDs).[24]

Transmission electron microscopy (TEM) reveals finer structural levels than optical microscopy and suggests that ancient sword makers were unintentional nanotechnologists because of the nanometer-scale patterns of carbon in the hard edge of the sword. This "self-assembled heterophase nanostructure" is responsible for the improved strength of ancient sword blades, which are formed by a quench hardening process that allows martensitic transformation and the redistribution of trapped interstitial carbon to occur.[1,20] However, people acquired this detailed understanding of the structure−property and processing−structure relationship long after the sword was invented. In other words, materials engineers used to develop materials for practical applications using trial-and-error approaches and matched each processing recipe with the ultimate properties without a clear picture of the hierarchical structure of materials.[25] As such, materials imaging has been frequently used to check whether the processing led to the desired micro-/nanostructure. Monitoring the grain size distributions and dimensional specifications in the steel industry and in semiconductor fabrications through the use of optical microscopy and critical dimension SEM (CD-SEM) are currently the main functions of materials imaging in industry.

Materials imaging also plays an important role in failure analysis to spot the root cause of a failure and identify what or who to blame the failure for. As such, forensic engineering and failure analysis are key to understanding the causes of various vehicle, marine vessel, and airplane accidents. However, it has rarely occurred to materials scientists or

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

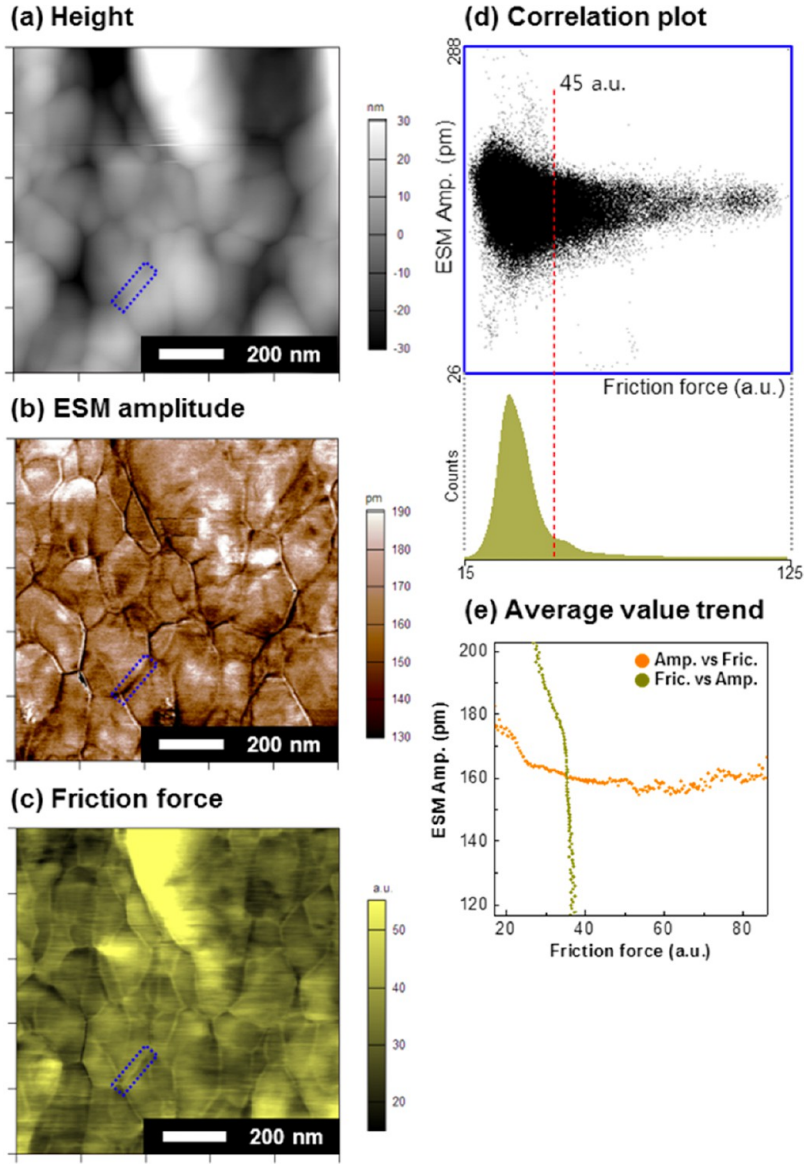


Figure 4. High-magnification AFM images of a composite anode (natural graphite (NG)/lithium silicon titanium phosphate (LSTP)/styrene−butadiene rubber (SBR)−carboxymethylcellulose (CMC)/Super-P). (a−c) Height, ESM amplitude, and frictional force images, respectively. (d) Pearson's correlation plot of the frictional force (*x*-axis) and electrochemical strain microscopy (ESM) amplitude (*y*-axis) and the histogram of the frictional force. (e) Trends of the average values of ESM amplitude as a function of the frictional force (orange) and the frictional force as a function of the ESM amplitude (olive). The ranges of the *x*- and *y*-axes are 99% of the full data. AFM images were acquired with a 4 V AC drive voltage to the AFM tip and a tip loading force of 400 nN. The blue lines in panels a−c are guides for the eye showing the valley between grains. Adapted with permission from ref 36. Copyright 2020 American Chemical Society.

engineers that materials imaging can radically accelerate materials design, discovery, and development.

ROLE OF MATERIALS IMAGING IN MATERIALS DISCOVERY AND DEVELOPMENT

Thus, it must be asked: What breakthrough in materials imaging created a paradigm shift in the discovery and development of materials? In answering this question, we will first discuss the evolution of materials imaging by studying an anecdote from Korean history. In the Chosun Dynasty, Jung Ho Kim was the person to correctly depicted the shape of the Korean Peninsula with detailed topographical information on a map.[26] According to an old document, he traveled around the country to measure the surface morphology using his own ruler and equipment. If this is a true story, he was the person in Korean history to become an AFM probe—or rather he was a physical stylus—that acquired the surface topography of the Korean peninsula at the meter scale. The map he completed in 1861 is called the "Daedongyeo Map" and is the most accurate map in Chosun Dynasty.[26,27] From this map, one can readily find the highest mountain in Korea, which is Mount Baekdu (see Figure 3). One can also see that the maps are stitched together to cover the entire peninsula, which is a practice similar to that of the stitching of AFM, SEM, and TEM images to cover large-scale areas to understand the hierarchical context of the nanoscale or even atomic-scale features.[28]

Currently, of course, one can use satellite and GPS information to obtain the map of the Korean Peninsula with height information and pick the same Mountain Baekdu as the highest mountain in Korea.[29] Moreover, we can stitch GPS information around the globe and pick Mount Everest as the highest mountain in the world within seconds.[30,31] This exercise provides us with insights into the importance of the accuracy, scalability, and speed of materials imaging to locate the peaks and troughs of the overall materials structural landscape in terms of space and time.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

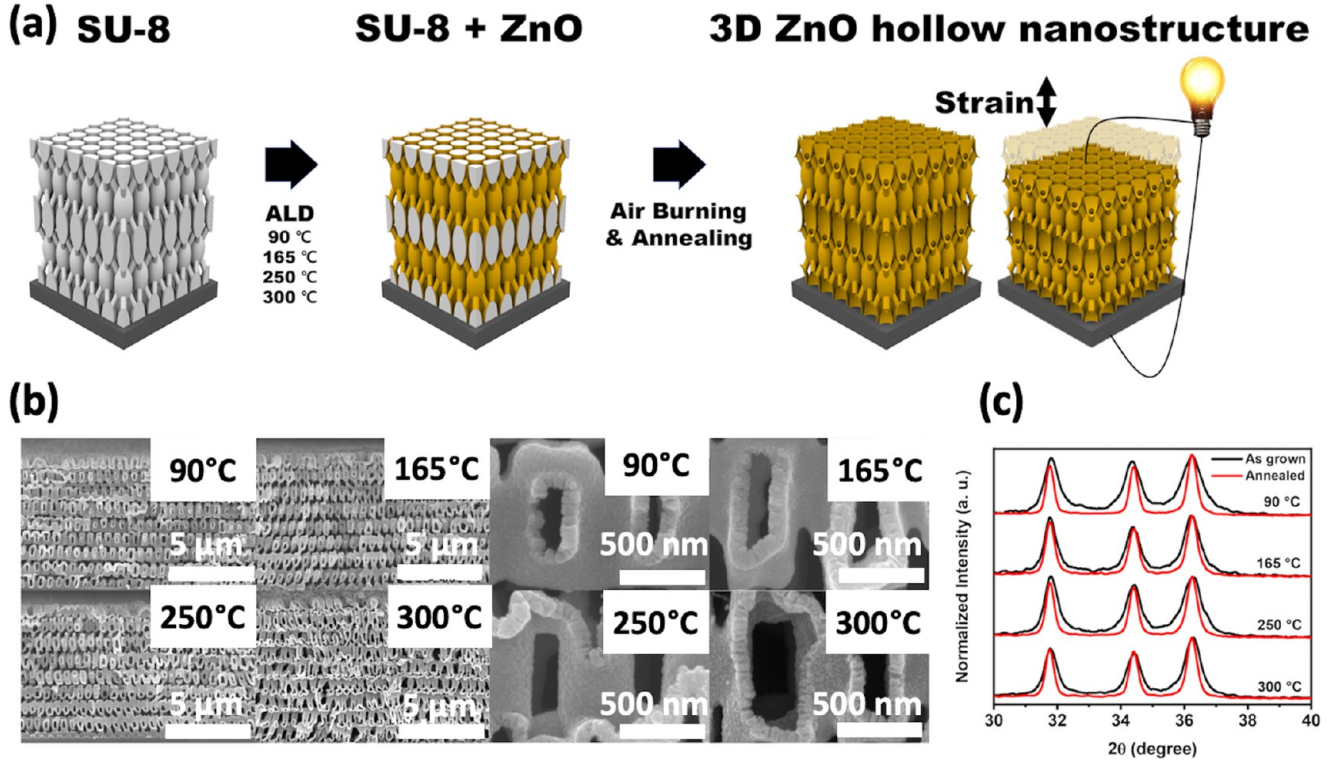


Figure 5. (a) Conceptual schematics and (b) SEM images of 3D-ZnO hollow nanostructures deposited at 90, 165, 250, and 300 °C after removing the epoxy template. (c) XRD patterns of the 3D-ZnO hollow nanostructure before (black) and after additional annealing (red) at each deposition temperature. Reprinted with permission from ref 37. Copyright 2020 Elsevier.

However, what use is an accurate multiscale geography map to lay people? One immediate use is to estimate the distance between two places where one wants to, *e.g.*, travel to visit their friends' house. Additionally, it may help people decide how long it will take to move from one place to the other and which route may be their best choice. However, without knowing the traffic situation of various roads connecting two places, the estimate will be rough at best. Having real-time traffic information layered on top of the accurate map will improve the accuracy of the travel time from one place to the other and help people choose the route with the shortest travel time. In materials science terminology, simultaneously knowing both the structures (road morphology) and properties (traffic status and/or rate of flow) in real time can help users determine the optimum route to tailor their functional performance (travel from place to place).

We show the incredible strength of imaging through the following example of electrochemical strain microscopy (ESM) to map surface morphology overlaid with the distribution of Li ionic motion (Figure 4).[32−36] ESM can simultaneously reveal both the surface grain structure and diffusion property of Li ions *in situ*, which helps to elucidate the routes taken by Li ions moving from one site to another. *A priori* knowledge that the ions move along but not perpendicular to the layers in the cathode materials allows us to better understand why there is a strong correlation between the orientation and electrochemical activity of each grain.

Furthermore, imagine you have a map of the world and you overlaid a map of a specific monetary value rubric as a function of time. You can make a strategic decision to convert sea to land or vice versa if the initial cost can be quickly recovered by a much higher revenue flowing into the region of interest. Thus, land reclamation or the construction of canals can be understood from this perspective.

If removing some of the materials at the nanoscale can enhance the material properties and hence the performance and value, one can then design better performing materials. A good example is the case of a ZnO nanotruss structure, in which both the piezoelectric coefficient and elastic limit drastically improve in comparison to those of bulk ZnO materials (see Figure 5).[37] Here the researchers in effect removed nanoscale parts of the structure in a regular manner using proximity field nanopatterning (PnP) and atomic layer deposition (ALD), which could extend the elastic limit of brittle materials because a large flaw is less likely to occur in a small structure, which in turn enhances the strength of the material, according to the Griffith criterion. In addition, they attributed the improved piezoelectric properties to less constraint on the structure.[37]

MULTISCALE STRUCTURE−PROPERTY IMAGING

Now we can answer the following question: What breakthrough in materials imaging created a paradigm shift in the discovery and development of materials?

We suggest that the development of various microscopy and diffraction tools with the ability to map the structure, property, and performance of materials at multiscale and in real time enabled people to think that materials imaging could radically accelerate materials discovery and development.[11,38,39]

One of the examples that shows the power of structure−property imaging at the nanoscale is the development of future materials for emerging nonvolatile memory devices.[40] For example, the thickness and area of the capacitor are two important structure parameters, which can be optimized on the basis of the ferroelectric domain imaging to maximize its property such as the rate of polarization switching.[40] Elucidating the polarization switching mechanism by identifying the rate-determining step, one can design the capacitor to be large-area−low-thickness film or small-area−high-thickness rod. Two of the key performance parameters of a memory device are the volume of the smallest possible unit of information (in other words, information storage density in terms of bit/inch^2) and the speed at which one can write and read this information bit.[41,42] Consider the design of a ferroelectric capacitor to store the information and write/read such information at high speed. In

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

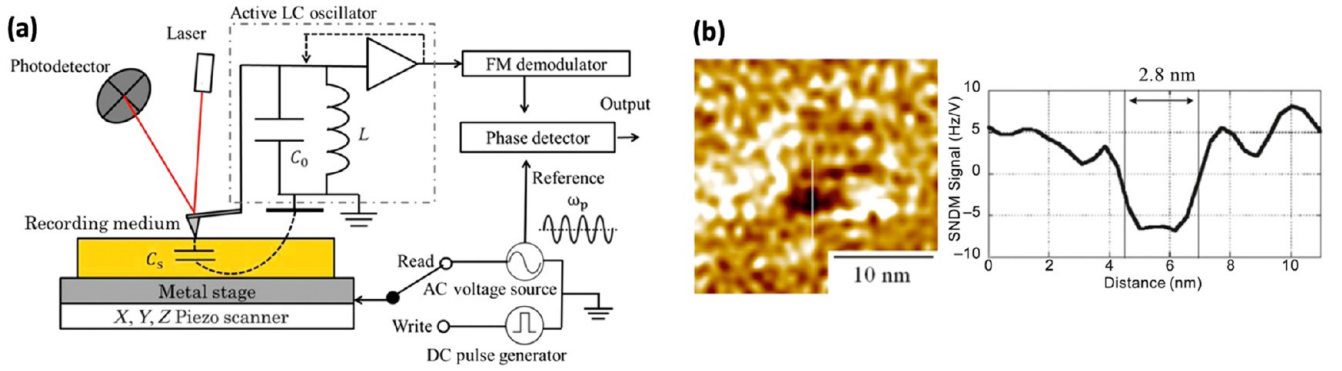


Figure 6. (a) Schematic diagram of the ferroelectric data-storage system based on scanning nonlinear dielectric microscopy (SNDM). (b) Smallest artificial nanodomain represented as single dots with a diameter of 2.8 nm, as observed by using SNDM. (Left) SNDM image and (right) profile taken along the white line in the left image. Note: C_0 is the built-in capacitance including the stray capacitance, L is the built-in inductance, ω_p is angular frequency of the applied voltage, C_s is the tip−sample capacitance, LC is the inductance−capacitance, AC is the alternating current, DC is the direct current, and FM is the frequency modulation. Reprinted with permission from ref 45. Copyright 2008 The Japan Society of Applied Physics.

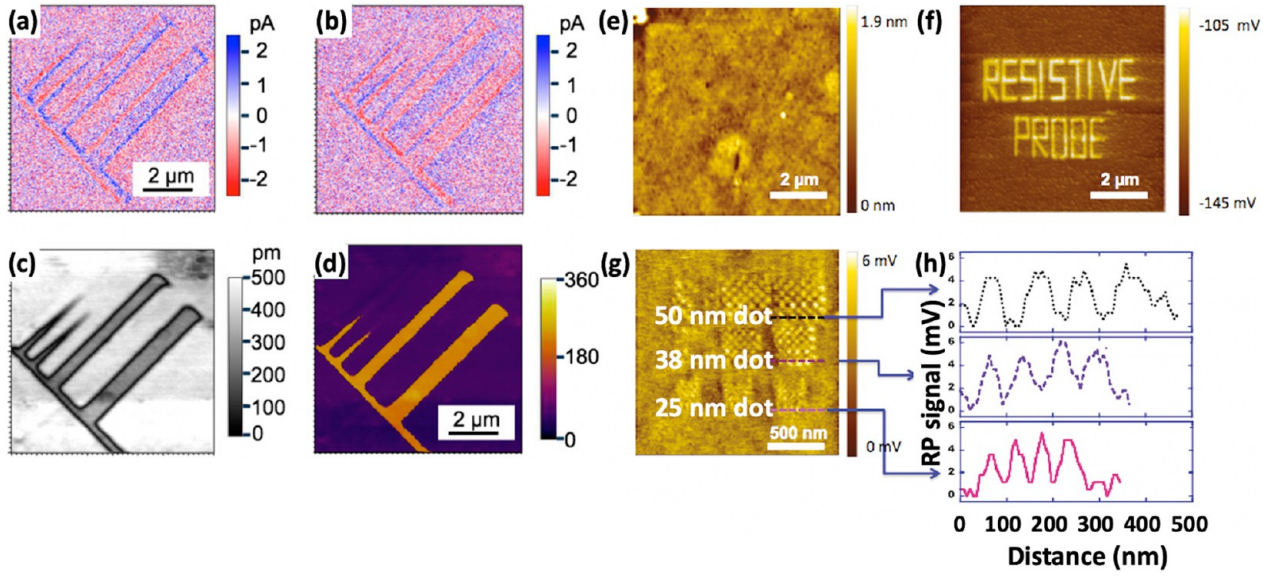


Figure 7. Trace and retrace CGM images of artificially decorated domains with different sizes. CGM images taken on ribbon-shaped domains poled by 6 V to the bottom electrode of 85 nm thick $LiTaO_3$ films with a (a) left-to-right scan (trace) and (b) right-to-left scan (retrace) using Pt tips at scan frequency of 40 Hz. Piezoresponse force microscopy (PFM) (c) amplitude and (d) phase images obtained in the same region at a scan frequency of 1 Hz where a bright-phase contrast corresponds to the positive (upward) domain and a dark-phase contrast corresponds to the negative (downward) domain. Reprinted from ref 59. Copyright 2014 The Authors under Creative Commons International 4.0 Attribution license (https://creativecommons.org/licenses/by/4.0/), published by The National Academy of Sciences (NAS). Manipulation and detection of surface charges using SRPM. (e, f) Topography (e) and SRPM (f) images of artificially polarized domains with letters showing "RESISTIVE PROBE". (g, h) Checkerboard patterns of alternating positive and negative domains imaged by SRPM (g) and the selected line profiles of domains with distances of 50, 38, and 25 nm (h). Reprinted with permission from ref 58. Copyright 2011 American Chemical Society.

addition to the performance parameters, you will also need to know the reliability parameter, which may be the lifetime of your ferroelectric capacitor.[43]

How can materials imaging help design the best ferroelectric capacitor? First, one may want to know the size limit of the ferroelectric domain that consists of the same polarization within the region of interest. Using synchrotron X-rays as the probe, the structure of materials can be imaged in a reciprocal space to map the ferroelectric polarization, thereby providing us the answer once we know the minimum thickness along with the minimum width of the ferroelectric domain at such a thickness. Fong and his colleagues at Argonne National Laboratory reported that a $PbTiO_3$ thin film can sustain its ferroelectricity down to three-unit cells,[44] and Cho's research group at Tohoku University[45,46] successfully imaged the smallest ferroelectric domain in $LiTaO_3$ thin films with diameters of less than 3 nm (see Figure 6).

Furthermore, Hong *et al.*[47,48] found the rate-determining speed of ferroelectric domain switching that could help engineers design the aspect ratio of the ferroelectric capacitors to maximize the switching speed. In addition, Colla *et al.*[49] reported the distribution of frozen domains that were responsible for polarization fatigue in ferroelectric capacitors as a function of switching cycles, which is directly related to the lifetime of the capacitor. Therefore, the simultaneous nanoscale mapping of the domain structures and properties (such as switchability) provided important insights into the design of fatigue-free ferroelectric capacitors.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

On a side note, the development of imaging tools, such as piezoresponse force microscopy,[50−56] scanning resistive probe microscopy,[57,58] or charge gradient microscopy,[59,60] can lead to the development of information storage devices or vice versa (see Figure 7).[61,62]

It is worth noting that there are pioneering groups promoting multiscale structure−property imaging in several places, such as the Integrated Imaging Institute (I3) at Argonne National Laboratory[63] and the Center for Nanophase Materials Sciences (CNMS) at Oak Ridge National Laboratory.[64]

MATERIALS BY DESIGN: HIGH-THROUGHPUT SCREENING USING DENSITY FUNCTIONAL THEORY AND MACHINE LEARNING

A prominent change in materials science during recent decades has been the active integration of computer simulations to find materials, predict properties, and understand phenomena. Using first-principles calculations, many research groups have proposed the idea of identifying compounds with interesting properties using computers and high-throughput screening algorithms.[2,65−68] Calculations, particularly those based on density functional theory (DFT) and other *ab initio* techniques, are more easily scaled across diverse chemical spaces than experiments.[67] For example, all of the elements and the possible crystal structures can be used to calculate material properties such as the conductivity, permittivity, polarization, magnetization, diffusion coefficient, elastic modulus, formation energy, and adsorption/desorption energy. However, since DFT suffers from the limitations, namely, the small number of cells (*e.g.*, 1000 atoms) and a temperature of zero Kelvin, while underestimating the band gap for certain classes of materials, materials scientists have started to combine it with conventional molecular dynamics, phase field modeling, and continuum physical modeling tools. Furthermore, to enhance the calculation efficiency, researchers have started to use machine learning (ML) to calculate the density of states and electronic band structure.[68]

Machine learning (ML), an algorithm that underpins a majority of artificial intelligence (AI),[69] is rooted from linear regression—of which algebraic procedure was established by Adrien-Marie Legendre in 1805[70]—and pioneered by Arthur Samuel in 1959 who defined it as the field of study that enables computers to learn without being explicitly programmed.[71] While ML models and algorithms have been developed since the 1950s,[72,73] it is only in the recent decade that the systematic generation and curation of data have enabled ML to nucleate frontiers across many fields including materials science and engineering.[66,74−78] ML is especially suitable for exploratory tasks that feature exponentially complex solutions, which is exemplified by the recent triumph of AlphaGo and AlphaGo Zero in solving the problem of Go, which has an estimated 10^{170} potential outcomes.[79,80]

The ability of ML to interpolate and in some cases extrapolate from a set of training data to explore unknown spaces makes it a fascinating solution to many challenges in materials science including the problem of materials discovery and development.[38,77] Currently, approximately 10^6 crystalline materials and 10^9 molecules have been explored and investigated computationally and/or experimentally. Nevertheless, this constitutes a tiny fraction of potential crystals and molecules in the universe, where the number of only small organic molecules is estimated to be 10^{60}.[68]

High-throughput DFT screening limits the search space to thousands of compounds.[65] Therefore, ML offers a solution to the limitation of the small search space of DFT by making predictions of materials or properties from existing data, which in turn can drive the generation of more data that can be used to further refine the ML models.[38] Recently, Noh *et al.*[81] reviewed the latest progress in machine-enabled inverse materials design, which can be categorized into high-throughput virtual screening, global optimization, and generative models. The inverse design has been performed by exploring the chemical space effectively toward the target region. However, the grand challenge of inverse design is the physical realization of predicted materials, which attests to the importance of developing an experimental feedback loop for discovered materials.[81]

The MGI has provided large, public databases of computed material properties.[3,66] Similar attempts have been made around the globe, such as the Materials Research by Information Integration Initiative in Japan,[4] Novel Materials Discovery in the E.U.,[5] Vom Materials Zur Innovation in Germany,[6] Materials Scientific Data Sharing Network in China,[7] and Creative Materials Discovery in Korea.[8]

Examples of general-purpose databases with high chemical diversity include the Materials Project,[66,82] Citrination platform,[83,84] Automated Interactive Infrastructure and Database for Computational Science (AiiDA) platform,[85] AFLOWLIB,[86,87] Open Quantum Materials Data (OQMD),[88,89] Novel Materials Discovery (NOMAD) repository,[90] JARVIS-DFT[91] (JARVIS[92]), Inorganic Crystal Structure Database (ICSD),[93] and Materials Data Facility (MDF)[94] to name a few. These resources contain millions of computational "measurements" of material properties, *e.g.*, the formation enthalpy, electronic band structure, and elastic moduli, which can be systematically searched. The ability to rapidly generate reliable material data in this manner improves every year as computing costs decrease; thus, the study of materials with theoretical methods becomes more accurate, and the software to apply these techniques becomes more powerful and more accessible to a larger audience.[95]

The Atomate library,[96] developed by the Materials Project collaboration, uses several underlying libraries to create sophisticated materials models, manage workflows on supercomputing centers, and provide error correction. Atomate implements many common materials workflows and was used to create the Materials Project database.[82]

The Materials Data Facility (MDF) operates two cloud-hosted services, data publication, and data discovery, with features to promote open data sharing, self-service data publication, and curation and encourage data reuse, layered with powerful data discovery tools.[94] In collaboration with Materials Data Facility at Argonne National Laboratory, Phatak *et al.* implemented an automated data curation workflow for the transmission electron microscope that imposes minimal burden on users for additional information, yet collects data in a form amenable to automated analysis and machine analysis.[97]

Nonetheless, ML based on DFT has its own limitations. They employ a particular exchange functional, and thus in some cases the data are not accurate. In addition, the predictions need to be validated by experiments. As such, the data quality and reliability have become the most important challenges. Data uncertainty can come from many sources, such as computational errors from unsatisfactory approximation, experimental errors, and the intentional omission of important parts due to confidentiality or the fear of being copied by the fast followers.

Many existing experimental data repositories are still either too small or too inconsistent (*e.g.*, different experimental conditions, measurement techniques, or different simulation input choices) for high-quality ML models.[98] In addition, most of the databases are commercial products requiring a license, and programmatic application program interfaces (APIs) for large-scale data access are rarely implemented.[68]

MULTIMODAL AND MULTISCALE MATERIALS IMAGING: RISE OF USER FACILITIES

One way to tackle the problem of a small amount of experimental data is to tap into the very large data set created by user facilities around the world such as Advanced Photon Source (APS) at Argonne National Laboratory, SLAC National Accelerator Laboratory, Spallation Neutron Source (SNS) at Oak Ridge National Laboratory, nanoscience centers at six national laboratories in the U.S., Spring-8 in Japan, European Synchrotron Radiation Facility (ESRF) in France, Diamond Light Source in the U.K., and Pohang Accelerator Laboratory in Korea. With the policy of open and transparent data publication, and more efforts toward the standardization of data sets created by these user facilities, this is the right time to use such experimental data sets to validate DFT-based ML models and create experimental data-based ML models.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

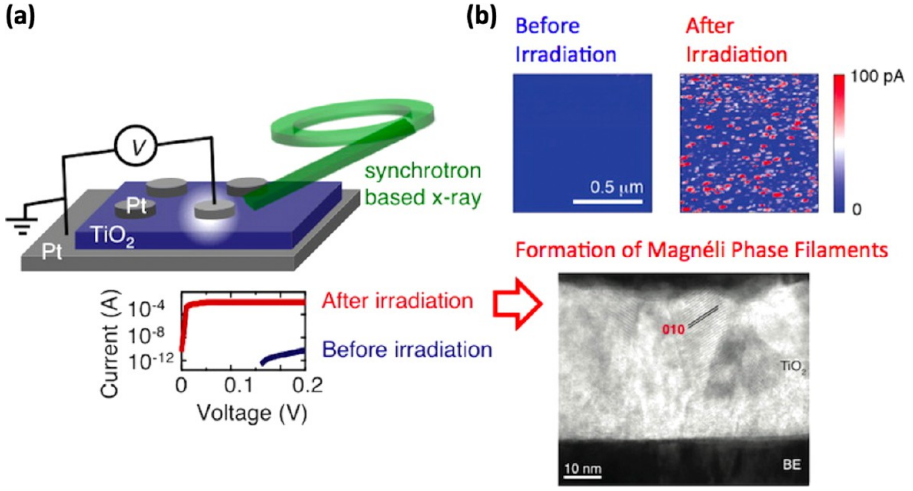


Figure 8. (a) Schematic showing the electrical measurement of a Pt/TiO_2/Pt cell with the synchrotron-based X-ray irradiation. *I*−*V* curves of the cell, measured before and after the prolonged irradiation. (b) C-AFM images of the pristine and the X-ray-irradiated surfaces. Two-dimensional current map measured at 0.02 V show distinct differences between the unexposed pristine and X-ray-irradiated regions. Conduction paths of 28 ± 14 nm diameter are clearly seen in the irradiated region. High-resolution electron microscopy (HREM) image of the filament-like region. Lattice spacing of d = 0.67 nm is clearly seen which corresponds to [010] of the Magnéli phase of Ti_4O_7. Reprinted with permission from ref 99. Copyright 2014 American Chemical Society.

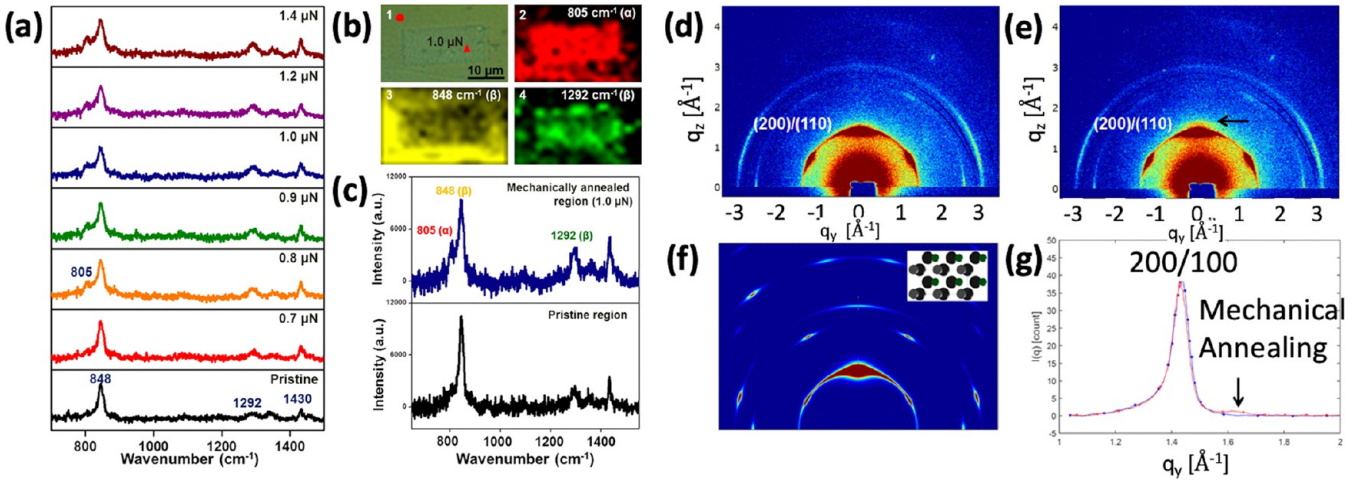


Figure 9. Raman spectra and mapping images of the P(VDF-TrFE) film before and after mechanical annealing. (a) Raman spectra of pristine and mechanically annealed regions as a function of mechanical force from 0.6 to 1.4 μN. (b) Optical microscopy image (b-1), Raman mapping images at 805 (b-2), 848 (b-3), and 1294 cm^{-1} (b-4), and (c) Raman spectra measured at the circle-marked position (pristine region) of image b-1 and at the triangle-marked position of image b-1 (mechanically annealed region). Grazing-incidence wide-angle X-ray scattering (GIWAXS) data of the (d) pristine region and (e) mechanically annealed region with a mechanical force of 1.0 μN, (f) simulated GIWAXS of the (200) orientation of the P(VDF-TrFE) film (inset: top view of the unit cell), and (g) azimuthally averaged data. The phase peak is denoted by the arrows in panels e and g. Reprinted with permission from ref 28. Copyright 2015 American Chemical Society.

While DFT-based ML has emerged as the central part of Materials by Design, with the imaging and characterization capabilities of user facilities, researchers have started to build a holistic picture of emerging phenomena in functional materials; thus, this process could potentially identify the use of existing materials in applications. For example, using multiple microscopy tools such as AFM and TEM, and synchrotron X-rays as a probe to change the structure of materials, Chang *et al.*[99] found that a photovoltaic-like effect can reversibly modulate the conductance of TiO_2 thin films by a few orders of magnitude in response to the intensity of impinging X-rays (see Figure 8). This method led to a reversible nonvolatile conductance change when the photovoltaic-like effect is combined with a local phase transition into a Magnéli phase. The discovery of a reversible resistance change induced by X-rays provides a methodology to initiate the electroforming process in TiO_2 thin films.

Choi *et al.*[28] discovered the mechanical annealing process, which could control nanoscale material properties and molecular orientation using intense local stress. They directly applied mechanical stress, which could induce irreversible plastic deformation, to a P(VDF-TrFE) thin film using a nanoscale tip at room temperature. The regions with plastic deformation did not show any significant damage or delamination. The vertical and lateral piezoresponse amplitudes measured after the mechanical annealing were 28% and 102.4% higher than those of the pristine film. In addition, randomly oriented lamellar crystals were well-aligned along the applied stress direction (see Figure 9). Therefore, they envisioned that mechanical annealing using intense local stress would be widely used to improve and control the local material properties in the polymeric thin films for fabricating high-performance piezoelectric devices.

Kalinin *et al.*[38] reviewed the challenges and opportunities for extending electrochemical characterization probes to the nanometer

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

and ultimately atomic scales (scanning probe and electron microscopy), and integration of spatial heterogeneity of local properties to understand macroscopic properties. In addition, they envisioned that the structure−property relationship on a single atomic-defect level could be unveiled by combining multiscale materials imaging with ML models.[100]

At this point we can answer the question we posed in the beginning: How can we discover and develop materials in response to the challenges of the fourth industrial revolution, the increase in national protectionism due to the COVID-19 era, and the rise of artificial intelligence, open data policies, and user facilities around the world?

With more powerful imaging tools that can map the structure and properties as a function of space, time, momentum, and energy dimensions, one can build a powerful library that correlates the complex relationship between structures and properties at different scales, which could be used to reverse engineer materials of interest with given performance specs.

MATERIALS AND MOLECULAR MODELING, IMAGING, INFORMATICS, AND INTEGRATION (M3I3) INITIATIVE

Recently, KAIST announced 10 flagship research fields, which include the KAIST Materials Revolution: Materials and Molecular Modeling, Imaging, Informatics and Integration (M3I3) (see Figure S1).[12] The M3I3 initiative aims at reducing the time for the discovery, design, and development of materials based on multiscale processing−structure−property relationships and materials hierarchy, which will be quantified and understood through combination of ML and scientific insight.

As such, M3I3 is an algorithm to perform a reverse engineering of future materials. Fast followers usually copy the products of the first movers by reverse engineering them. For example, in the case of state-of-the-art battery products, competitors dissect them into pieces and analyze the structure and composition of each part, such as the cathode, anode, electrolyte, and separator. This so-called "reverse engineering" is the most inexpensive way to catch up with the forefront runners in the ever-expanding competitive world. Front runners also need a way to defend themselves and aggressively keep distance from their competitors, which is why they invest a huge amount of resources into the research and development of materials, devices, systems, and platforms, and file patents all over the world. M3I3 provides a means to effectively achieve this goal by mimicking a "reverse engineering" strategy with a higher level of creativity. Thus, M3I3 can reverse-engineer future materials of interest with superior performance and reliability as well as with minimal cost and environmental impact.

Reverse engineering starts by analyzing the structure and composition of the cutting-edge materials or products. Once we determine the performance of our targeted future materials, we need to know the candidate structures and compositions for producing the future materials. This knowledge can only be available if we know the structure−property or the property−structure relationships of all materials and molecules at all scales. High-quality multiscale and multidimensional experimental data will be the key to the success of our approach.

APPLICATION OF M3I3 TO BATTERY MATERIALS

Among many types of materials, we aimed to apply M3I3 to rechargeable battery materials. The reason we chose battery materials was because of the relatively long history of materials development, initiation of the MGI, and intuitive and qualitative understanding of multiscale physical/chemical/electrical properties.[101−103] Since Sony Co. commercialized the lithium-ion battery (LIB) in the 1990s, the demand for rechargeable batteries has become so pervasive that the capacity requirement is now common knowledge to most of the people. With the consensus and recognition of climate change caused by greenhouse gases, governments around the world have launched numerous initiatives on renewable energy technologies and electric vehicles to tackle the climate challenge.[102]

Researchers around the globe added at least 202,756 publications on batteries from 2010 to 2020, representing a 320% growth in the total number of journal papers based on the Web of Science search with query of "batteries".[102] In the meantime, with competition between major car companies such as Toyota, GM, BMW, Audi, Mercedes-Benz, Hyundai, and Tesla, the market for Li-ion batteries has surged with a market size of 36.35 billion USD in 2019 and is projected to reach 115.98 billion USD by 2027, growing at a compound annual growth rate (CAGR) of 15.6% from 2020 to 2027.[104]

Even though the battery market has undergone an exponential growth, the ultimate goals of research have remained the same: to increase the volume and weight energy density, increase charging/discharging rate, enhance lifetime (cycling durability), ensure safety, and minimize cost and environmental impact.

Energy density has steadily increased but still needs to increase to create a larger impact in cars, ships, and airplanes as well as grid storage applications. If we think of electrodes as a parking lot with cars, then a building will be the materials to maintain the structural integrity of the electrode whereas the cars will be the Li ions that drive in and out of the parking lot when the batteries are being charged or discharged. Although this metaphor applies mostly to the intercalation type electrode materials, it could also be applied to other future battery materials used for, *e.g.*, lithium−sulfur and lithium−air batteries if we add one more step of car transforming into a truck with a volume expansion. In such cases, we need to consider the design of parking space for trucks like Li_2S and Li_2O_2 that are the discharging products of lithium−sulfur and lithium−air batteries, respectively.[105,106] In addition, we should consider the reversibility between car and truck as well as the structural resilience of the parking lot against the volume expansion. Improved cyclability using mesoporous 3D conducting scaffolds as the cathode underscores the need for optimum porous electrodes to confine Li_2S and Li_2O_2, which could be an example of a resilient parking lot.[105−108]

MULTISCALE DESIGN OF MATERIALS STRUCTURE

Then the question arises: How do we design the filler materials to maintain structural integrity and empty space for Li ions to diffuse in and out of the electrode? At the atomic scale, one must think of the space being able to embrace the case when Li ions are present and the case when Li ions are absent.[103] From an electric force point of view, the filler atom should be able to change its valence freely to accommodate both the presence and absence of Li ions. Therefore, transition metal ions are the most likely candidates in our electrode materials.

Now, if we increase the scale to the lattice scale (0.5−1 nm), then we should consider the quantum mechanical effect, such as repulsive force between atoms or ions, due to the uncertainty principle.[109] As such, point defects such as vacancies and interstitial sites should be considered altogether. At this scale, inorganic materials form typical structures such as perovskite, spinel, or layered structures that can host Li ions easily without destabilizing the structural integrity when charging and discharging occur.[103] Notably, the typically observed sites are either octahedral or tetrahedral.

As we scale up to grain and grain boundaries, the length scale would be between 100 nm and 10 μm, where the orientation of grains and their nature, the width and length of grain boundaries, and the characteristics and density of grain boundary junctions start to play important roles. Will grain boundaries act as fast diffusion paths of Li ions or as additional parking lots for Li ions? Will they help maintain the structural integrity during the volume changes that accompany with charging and discharging?

If we approach the scale of hundreds of micrometers to tens of millimeters, then we start to see composite materials comprised of inorganic particles, polymeric binders, conductive additives, and electrolytes. The shape of particles, the stress/strain gradient in radial and circumferential directions, the composition gradient (core/shell), the interface between particles and binder, the network structure, and the composition of additives will start to play major roles in the energy density of electrode materials.

To control the properties of materials in regard to a performance point of view, knowing all of the details of the multiscale structure−property relationship can be a daunting job. As such, many practical engineers rely on direct processing−property or processing−perform-

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

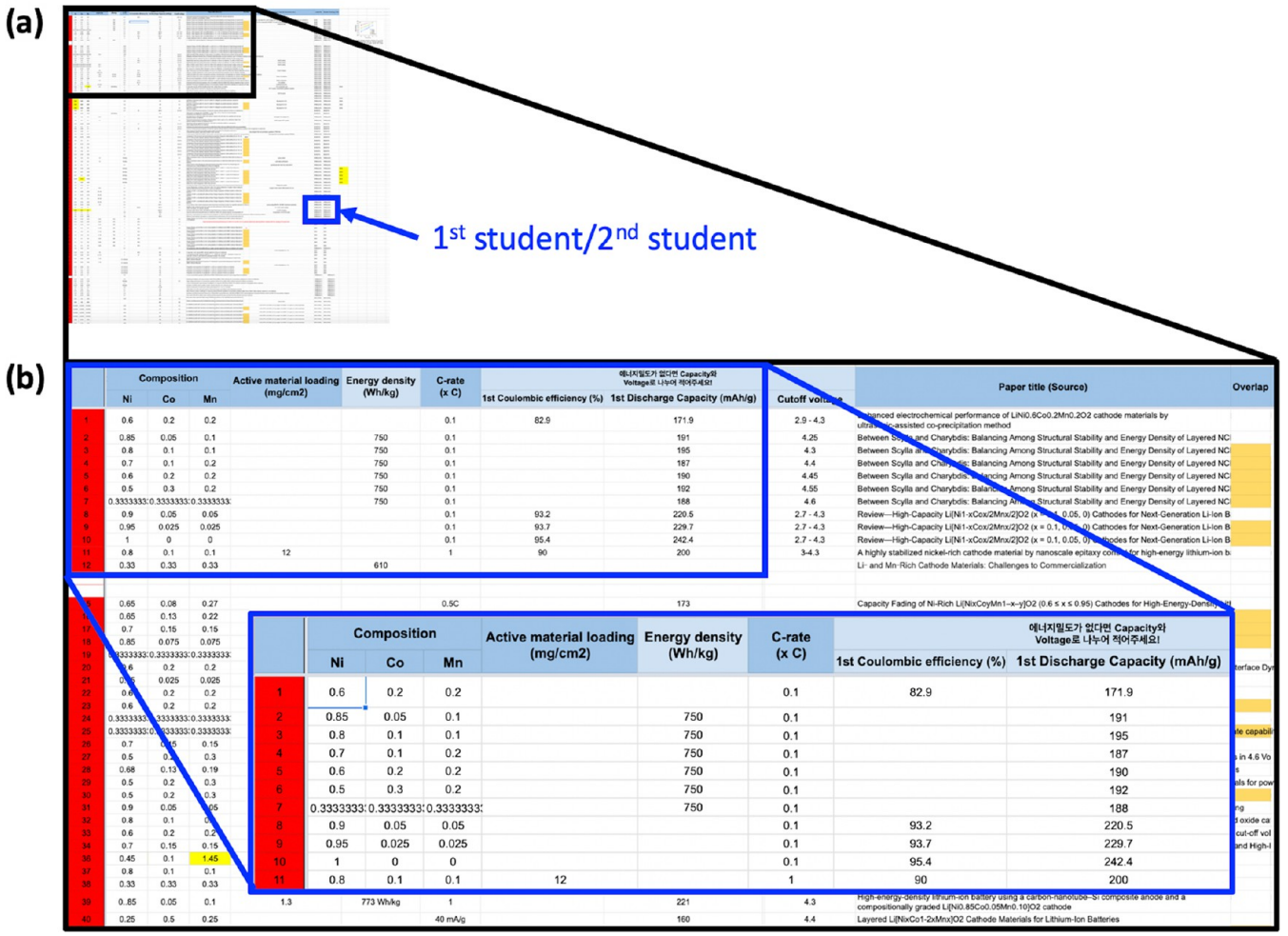


Figure 10. (a) Part of the M3I3 data sheet extracted from journal papers on NCM cathode materials from 2004 to 2019. (b) Magnified portion of the data sheet marked by the black box and the blue box inside the black box (inset). A full list can be found in ref 111.

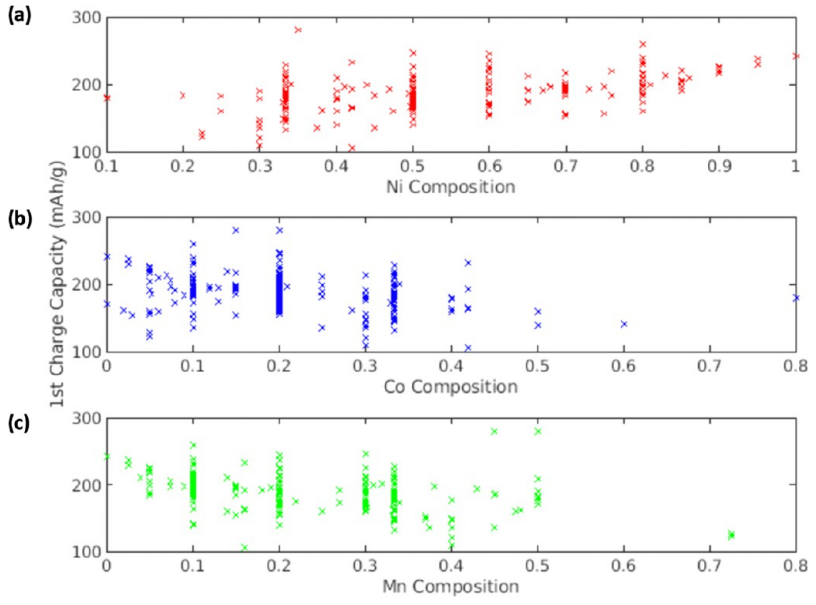


Figure 11. First charge capacity (mAh/g) *vs*: (a) Ni, (b) Co, and (c) Mn compositions. This plot was created by MATLAB.

ance relationships and optimize the properties and performance within the boundary conditions of the processing equipment and starting raw materials they have. Because most of the experimentalists tend to focus on so-called positive results, large amounts of unreported "dark data" are discarded, which we need to keep to truly find the global optimum point for maximizing the property and performance of interest.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

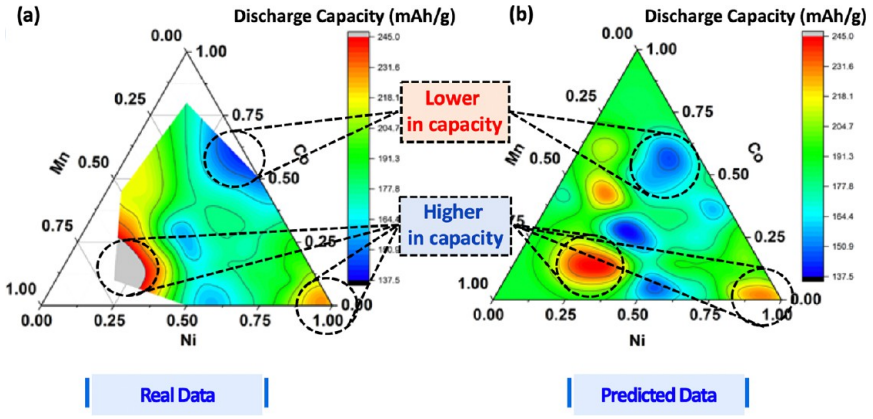


Figure 12. Capacity contour triangle plot as a function of the Ni, Co, and Mn compositions. (a) Left plot from the data extracted from the data sheet and (b) right plot based on the machine learning model developed using the same data sheet.

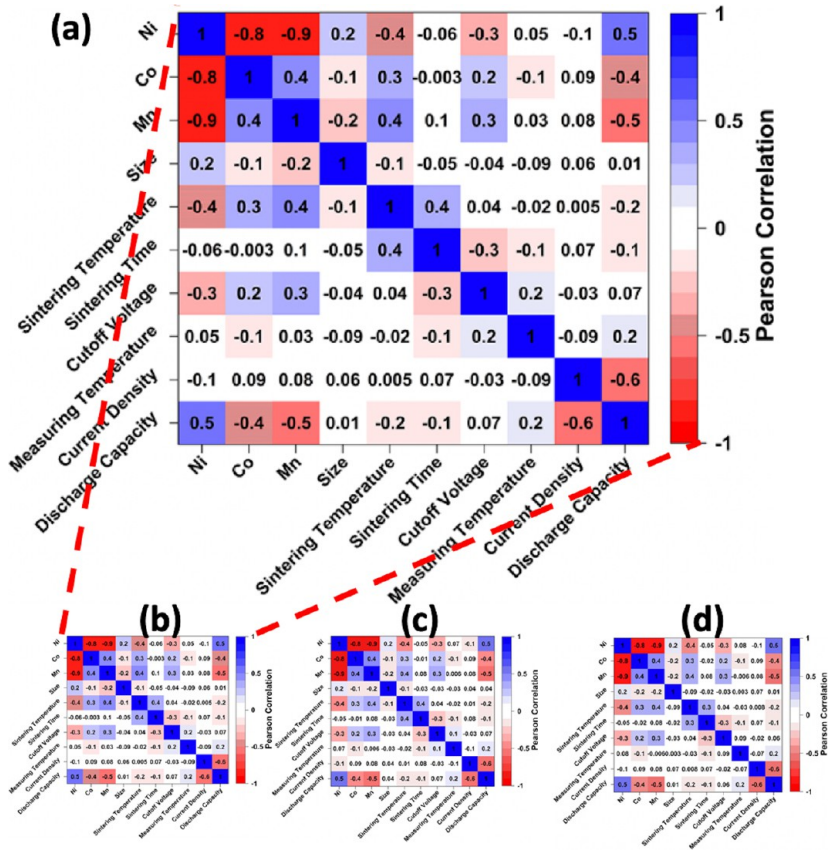


Figure 13. Pearson correlation coefficient between each parameter derived by (a, b) random forest (RF), (c) *k* nearest neighbor (KNN), and (d) MICE.

MATERIALS INFORMATICS AND MACHINE LEARNING FOR MATERIALS IMAGING AND INTEGRATION

Here we propose two independent but related approaches for tackling this issue within the M3I3 philosophy stated above.

One way is to use data mining from existing literature (mainly journal papers). The ideal method would be to use natural language processing (NLP) to create large and diverse material data sets.[110] However, since NLP is at the development stage, we used collective intelligence (in other words, 55 students worked together to fill out a data sheet reading the papers allocated to them) and created a data sheet where we identified approximately 900 important papers related to Li-ion batteries using NMC electrodes ($LiCoO_2$-based materials).[111] To minimize human errors, we assigned two students for each paper to double check the numbers input in the data sheet (see Figure 10). One type of error we observed during the analysis were the typos in the literature where the data in the table had an inconsistent composition.

Then, we used the composition of the NMC electrode as the main input parameter and energy density as the main output parameter. However, in doing so, we found that many of the papers lacked the information about the energy density and rather reported the output charge/discharge capacity as the proxy of energy density. As such, we changed our strategy to use the charge/discharge capacity as the output parameter.

Using a very simple linear regression and gradient descent technique, we were able to plot the discharge capacity as a function of the composition of Ni, Co, and Mn. As we confined our search within the NMC system, the total amount of Ni, Co, and Mn was assumed to be

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

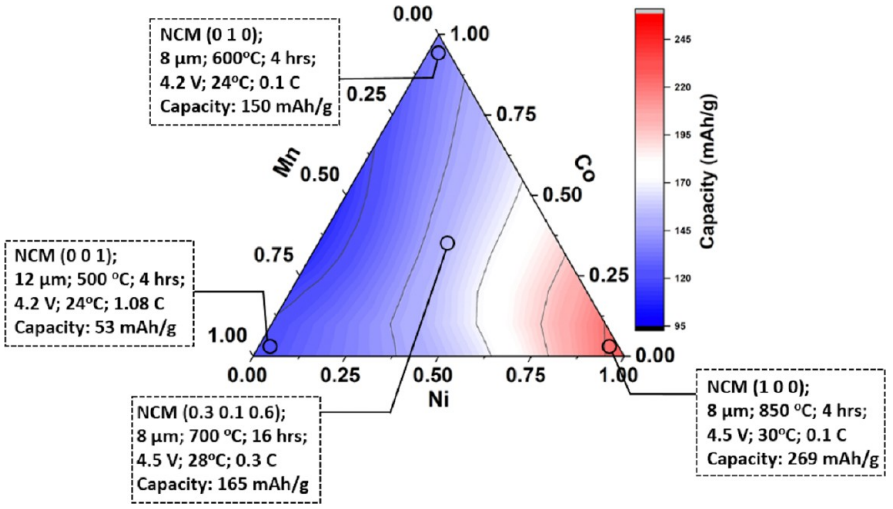


Figure 14. Capacity contour triangle plot as functions of composition (Ni, Co, and Mn), particle size, sintering temperature/time, measurement temperature, cutoff voltage, and C-rate.

Table 1. Experimental Validation of the Machine Learning Model Based on the Data Sheet in Reference 111 with $R_{train}^2 = 0.95$ and $R_{test}^2 = 0.85$

	composition											
no.	Ni	Co	Mn	particle size (μm)	sintering temp (°C)	sintering time (h)	cutoff voltage (V)	measuring temp (°C)	C-rate	actual capacity (mAh/g)	calculated capacity (mAh/g)	error (%)
1	0.5	0.2	0.3	15			4.5	25	0.1	190.0	191.3	0.7
2	0.5	0.2	0.3	15	900	12	4.5	25	0.1	181.1	180.5	0.3
3	1	0	0	5	650	12	4.5	25	0.1	214.0	236.3	10.4
4	0.95	0	0.05	5	650	12	4.5	25	0.1	214.0	231.3	8.1

constant, and therefore, we would have two degrees of freedom to change the composition of the NMC electrode. In other words, if we knew, *e.g.*, the Ni and Co compositions, the Mn composition could be obtained from the following relationship, [Mn] = 1 − ([Ni] + [Co]).

We used the code provided by the machine learning course in coursera.org,[69] and plotted the output discharge capacity as a function of the Ni, Co, and Mn composition, as shown in Figure 11. One can find the expected trend of the discharge capacity increasing as the Ni content increases. Another thing one can find is the popular composition of NMCs, where there are concentrated data points. The other thing is that there are other factors that contribute to the discharge capacity because at certain Ni compositions, the vertical spread is very large. Some other outliers indicate either experimental error or true innovation and therefore need special attention.

We found that these outliers include a sophisticated control of morphology, such as core/shell structures, concentration gradient particle structures, additive dopants, or Li-rich components. With these data included and using support vector machine (SVM), we built a triangle-shaped composition−first discharge capacity diagram, as shown in Figure 12. One can immediately spot two peaks in the predicted diagram and pick either Li-rich electrodes or Ni-rich electrodes as promising future electrodes for high energy density.

To learn the role of composition in the cathode without additional effects from multiple dopants, we restricted our search to a constant amount of Ni, Co, and Mn and decided to exclude Li-rich electrodes and solely focus on NMC electrodes to determine whether we could obtain the overall trend of moving to the Ni-rich composition, which is currently the market trend.

Although Li-rich cathodes ($Li_{1+x}Mn_{0.5+y}Ni_zCo_wO_2$, where $x + y + z + w$ is 0.5, also shown in Table S2) can reach higher output values of discharge capacity than Ni-rich cathodes, they often suffer from critical capacity fading and discharge voltage decay for prolonged cycles larger than 4.5 V.[112] The additional consideration of Li in cathode material complicates the overall thermodynamic stability of the cathode materials.[112] As such, Li-rich data from reliable resources is significantly smaller than that of the conventional NCM one. Therefore, our current approach focuses on simpler study of NCM. However, we intend to expand our approach to find the optimum Li-rich cathodes using transfer learning based on NCM.

In addition, we also wanted to know if other parameters reported in the literature affected the output discharge capacity as much as the reported composition of Ni, Co, and Mn. As such, we checked the correlation between the discharge capacity and sintering temperature and time, NCM particle size, cutoff voltage, measurement temperature, and current density (or C-rate). Figure 13 shows the Pearson correlation coefficient map between each parameter, where the NMC composition and current density (C-rate) have the largest influence on the discharge capacity. On the basis of this correlation, we employed three different supervised nonlinear imputation techniques (k nearest neighbor (KNN), random forest, and multivariate imputation by chained equations (MICE)), which statistically predict the missing values in the data sheet.

With KNN, the algorithm can find the missing data based on its closest k neighbors. While in random forest, the missing data are imputed on the basis of the nonlinear interaction within the variable. Finally, with MICE, the missing data are imputed on the basis of a conditional model of variable distribution. All three different methods operate in multidimensional space rather than a single imputation. With this complete data set as input features, the trained model is ready to be used to calculate the output voltage and charge capacity. Finally, we demonstrated the output capacity calculation based on approximately 9 million data sets using the trained model. This allows us to quickly map the output voltage or charge capacity on the basis of the interested compound ratio and their optimal operating conditions (see Figure 14). This calculation has been verified by experimental results with a maximum error of 10% deviation, as shown in Table 1.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

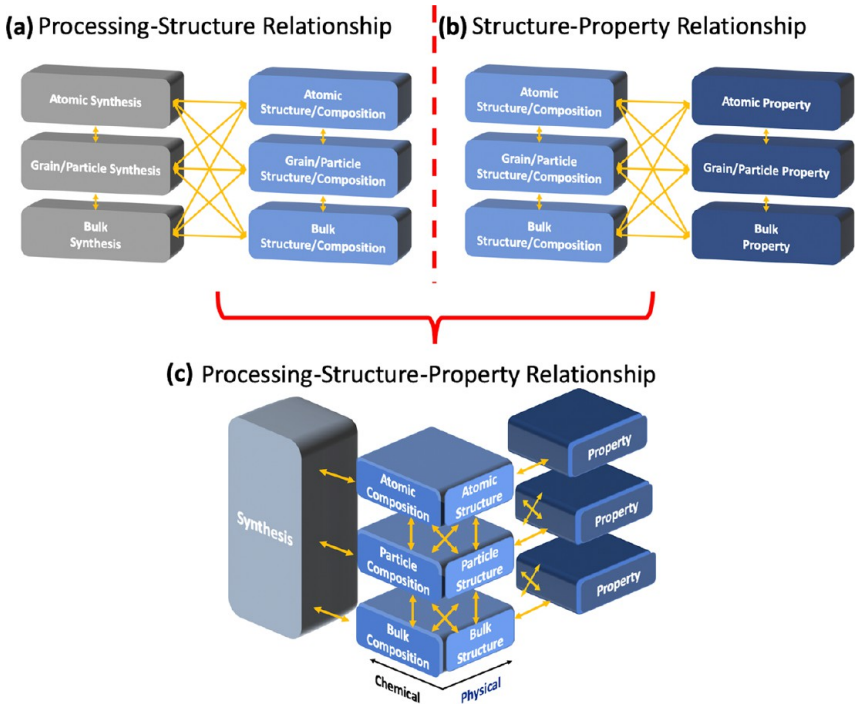


Figure 15. Two-step strategy of M3I3. Separate libraries of multiscale (a) "processing−structure" and (b) "structure−property" relationships. (c) Merging of these two libraries to create a seamless multiscale "processing−structure−property" library.

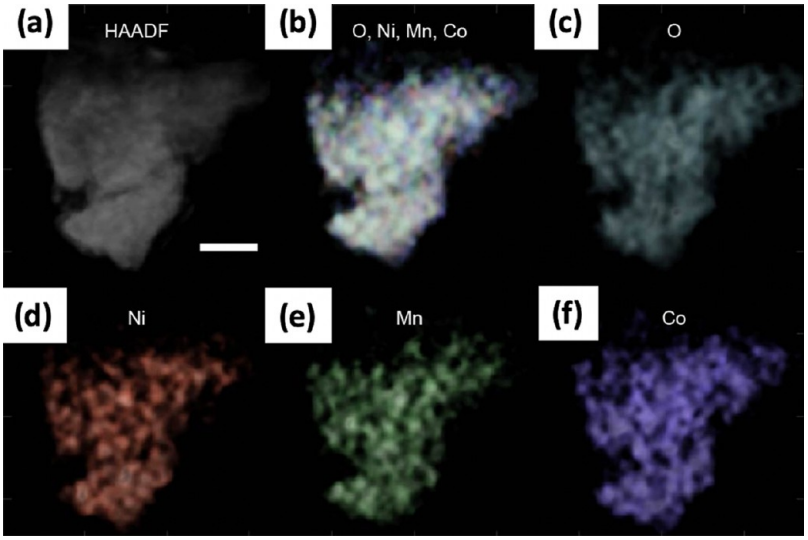


Figure 16. 0° projections of (a) HAADF and (b−f) STEM-EDX 3D tomography reconstructions for a layered $LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$ cathode material. Full 3D tomograms are provided in Supporting Information Video S1. Scale bar, 50 nm.

INTEGRATION OF STRUCTURE−PROPERTY AND PROCESSING−STRUCTURE RELATIONSHIPS

The other strategy is to use ML combined with physics- and chemistry-based models to independently correlate multiscale structure and property parameters and processing and structure parameters and finally merge those two to have a streamlined processing−structure−property model where both forward and inverse propagations are allowed (see Figure 15).

We first focus on the exploration of the structure−property relationship based on materials imaging as a quantitative tool. Recent progress in high-resolution, real-space imaging techniques such as scanning probe microscopy (SPM) or AFM,[113,114] TEM,[115,116] scanning transmission electron microscopy (STEM),[117,118] scanning tunneling microscopy (STM),[119,120] and atom probe tomography (APT)[121] has allowed the direct and efficient imaging of atomic columns and surface/interfacial atomic structures. These techniques enabled the direct visualization of the structure of materials, providing information on structural motifs that underlie crystals, grains, grain boundaries, dislocation cores, and quasicrystals, thereby leading to a fundamental understanding of the chemistry in these systems.

The spatial and temporal resolution of these methods have improved sufficiently to quantify the picometer-level displacement of atoms from original positions, thereby providing direct information about, *e.g.*, the ferroelectric polarization (or dipole moment), octahedral tilts and rotations, and electrochemical

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

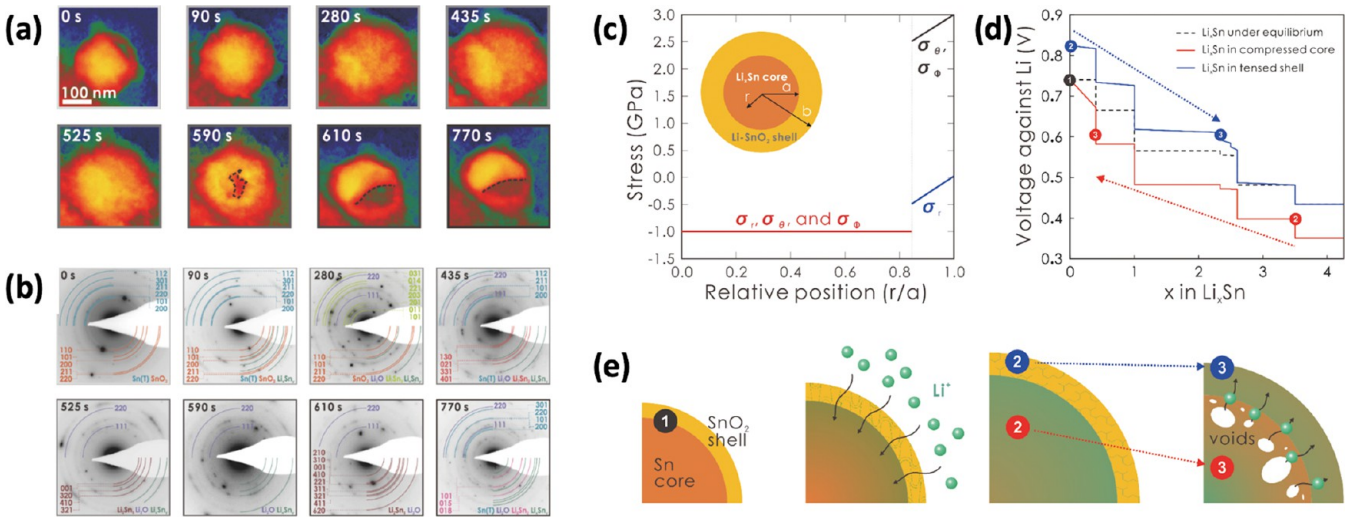


Figure 17. (a) Visualization of phase evolution in the Li−Sn alloy under the stress. Lithiation-induced volume changes during *in situ* lithiation of a Sn−SnO_2 core−shell particle. (b) Electron diffraction patterns taken at the respective moments. (c−e) Thermodynamic rationale for the strong stress−composition coupling: (c) stress distribution within the particle during lithiation according to the relative radial position from the particle center, with the inset showing the lithiated particle geometry; (d, e) stress effect on the equilibrium potential of the Li−Sn alloy system (d) and schematic illustrating the spontaneous core dealloying based on the stress-driven potential differences (e). The numbered circles indicate the discrete time frames during lithiation. Reprinted with permission from ref 131. Copyright 2019 The Authors under Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0/), published by Springer Nature.

strains. In addition, high-resolution STM and AFM can provide real-space atomic and electronic structures of material surfaces, visualizing structures of molecular vibration and lattice vibration (phonons), strongly or weakly correlated electrons, and chemical bonds (bond length and angle).[38]

The precise determination of the internal distribution of chemical elements within materials plays a crucial role in tailoring functional materials for desired properties. Especially for transition metal-based Li-ion battery cathode materials for which the structure−composition−property relationships basically govern their performance, accurate determination and control over the 3D distribution of different transition metals within the materials are essential.[103] It has been recently demonstrated that the segregation and migration of chemical elements (especially Ni) can substantially affect the surface reconstruction and degradation of the cathode during the battery cycling.[122,123] However, there are few techniques that can nondestructively map the 3D distribution of chemical compositions at the nanometer scale. Unlike many techniques that are either restricted to lower dimensions (2D projections or surface) or destructive (cannot detect the migration dynamics during the battery cycling), scanning transmission electron microscopy-based energy dispersive X-ray spectroscopy (STEM-EDX) tomography can safely provide the desired information. By combining the elemental specificity of STEM-EDX with tilt-series measurements and tomographic reconstructions,[124] the 3D distribution of individual chemical elements can be mapped at the nanometer scale.[122,125−127] Figure 16 shows the 3D tomographic reconstructions of high-angle annular dark-field (HAADF) tomography as well as the elemental tomography maps obtained from STEM-EDX tomography. HAADF tomography can reveal the overall 3D morphologies and internal density, while STEM-EDX tomography shows the 3D composition map. Since the STEM-EDX technique is nondestructive, it can be used to track the same specimen over the course of battery cycling, illuminating the dynamics of the migration of each chemical element during cycling.

Scanning probe and electron microscopies enable a broad range of spectroscopies, providing information about local electronic, mechanical, dielectric, and chemical properties. This local property imaging results in multidimensional data sets, which offer an opportunity to explore structure−property relationships at the level of single atoms and chemical bonds, linking local properties to local bond lengths and angles; furthermore, these techniques can be used at the level of micrometer-sized grains and particles, linking micrometer-scale properties (which for some cases represent the bulk properties) to local grains, grain boundaries, and mixed components in particles and the surrounding matrix.

For example, Kim *et al.* developed a visualization method that determined the distribution of components in battery electrodes using AFM.[36] They explored the dependence of ESM amplitude/phase and lateral force microscopy frictional force on the AC drive voltage and the tip loading force, respectively, and used their sensitivities as the markers for each component in the composite anode. This method allowed for the direct multiscale observation of the composite electrode under ambient conditions, distinguishing various components and simultaneously measuring their properties.

Electrochemical AFM visualizes the morphological evolution of electrodes and/or discharge products during the operation of battery model cells. The best benefit of electrochemical AFM among many *in situ* imaging systems is to preserve realistic cell conditions, *i.e.*, the temperature, pressure, liquid environment, and lack of external stimuli perturbation. This *in situ* observation has been used for lithium−oxygen cells to reveal the discharging and charging process. The nucleation and growth process of lithium peroxide (Li_2O_2) as a discharge product has been witnessed on highly oriented pyrolytic graphite.[128−130] The structure of Li_2O_2 is dependent on the electrolyte solution, inclusion of trace amounts of water, and additives. Electrochemical AFM imaging shows the growth of toroidal Li_2O_2 with

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

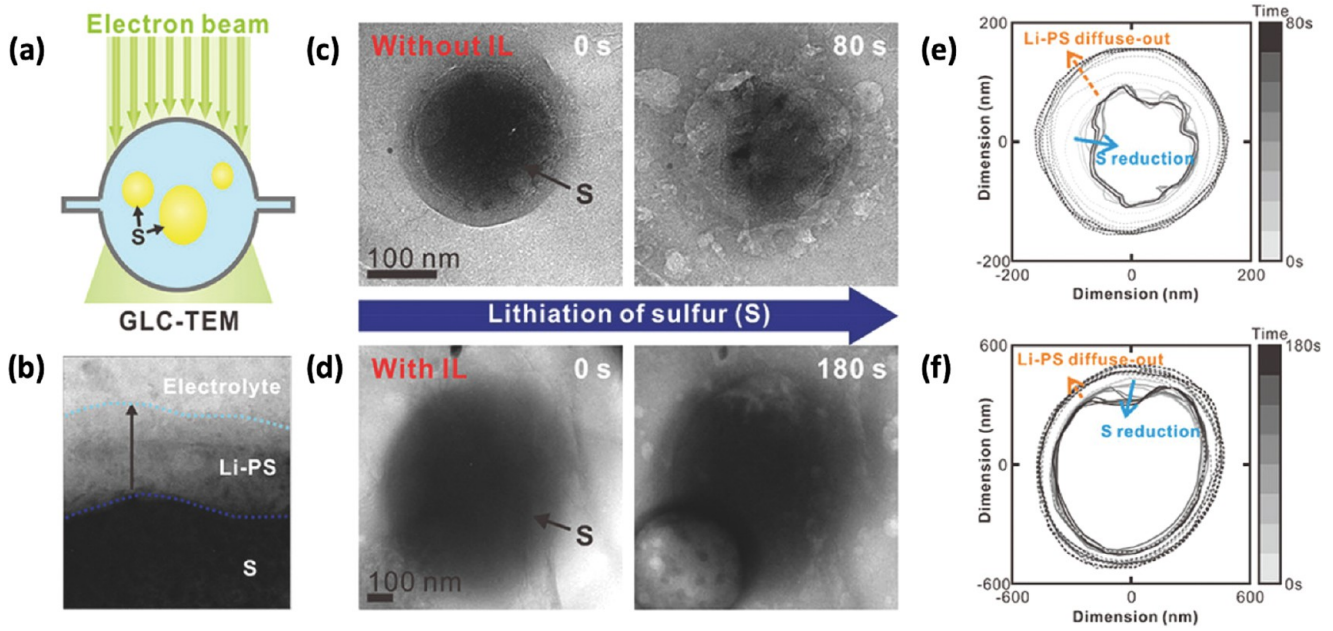


Figure 18. (a) Morphological and phase evolutions of sulfur nanoparticles during lithiation visualized by using *in situ* graphene liquid cell electron microscopy. (b) TEM image showing the interface between sulfur/Li-PS and the Li-PS/electrolyte. (c, d) Time-series bright-field TEM images showing the morphological evolutions of sulfur nanoparticles in (c) the IL-free and (d) the IL-containing electrolyte solution, respectively. The morphological evolution of sulfur and dissolved area of Li-PS are suppressed with the IL-containing electrolyte, while those with the IL-free electrolyte change significantly with lithiation time. (e, f) Outer boundaries of the sulfur core and Li-PS shown in panels c and d labeled with time, respectively. Reprinted with permission from ref 132. Copyright 2020 American Chemical Society.

DMSO,[130] while the growth of conformal Li_2O_2 is observed in a water-free tetraglyme electrolyte solution.[128,129] In addition, the decomposition of insulating Li_2O_2 is probed, and its morphological changes in response to charging potential are observed; results suggest the contribution of possible Li^+ dissolution from Li_2O_2 for triggering decomposition.[130]

As another example (see Figure 17), Seo *et al.* observed the evolution of structures and phases on Sn/SnO_2 core−shells during their lithiation using *in situ* graphene liquid cell electron microscopy (GLC-EM). They derived compositional information from the phases and stress information from the volumes of particles. These results offered the stress−composition coupling through the visualization of particle structures.[131]

Apparently, Seo *et al.* also directly visualized the formation and diffusion of lithium polysulfides (Li-PS) (see Figure 18). They observed the morphological and phase evolutions of sulfur nanoparticles during their lithiation and found that Li-PS formation was slowed in an ionic liquid (IL)-containing electrolyte. This result provides the structure and solubility relationships.[132]

However, with the growing complexity and closed-source nature of modern microscopes, the much-lauded promise of artificial intelligence (AI), and machine learning (ML) to revolutionize TEM experiment design, execution and analysis have not yet been fully realized.[133] On the other hand, X-ray crystallography have adopted open, standardized methods and data exchanges with enormous success. Automated X-ray experimentation is routinely conducted at scale, aided by easily accessible libraries of past work to plan and interpret future studies.[134]

These large volumetric data sets coming from automated experiments are nearly impossible to analyze without the help of computational tools. Image segmentation (the act of partitioning an image into multiple useful sections) can be the most difficult and time-consuming part of the data set analysis workflow. Having a human manually trace each image would take a very long time of continuous work.[135] In addition, complex images may contain a variety of artifacts such as color gradients, streaks, and foreign particles that are difficult for conventional contrast-based computational techniques to segment.

Machine learning methods have emerged as next-generation tools for segmentation of large data sets. Convolutional neural networks (CNNs) have been used for identification of dendritic patterns,[136] classification of steel microstructures,[137] segmentation of precipitates and nanoparticles,[138,139] phase mapping in multicomponent alloys,[140] classification of ambiguous microstructures,[141] denoizing of synchrotron X-ray computed tomography (XCT) experiments,[142] and calibrating the rotation axis in XCT.[143]

While, in principle, more data are a positive development, our ability to process and extract physical and chemical meaning from ballooning data sets has not kept pace.[133] In this regard, AI and ML methods trained on established physical and chemical models enable us to identify statistically significant features in large, noisy, and potentially incomplete data streams to build structure−property libraries, which leads to knowledge discovery.

Once a library of multiscale structure−property relationships is available as partially demonstrated above, the next step is to build a library of multiscale processing−structure relationships.[144] We will develop this library for Li-ion battery materials with a focus on the cathode.

The first task is to understand the current boundary conditions in terms of processing techniques used in the field of battery materials. We confined the options of synthesis to hydrothermal[145] and solid-state synthesis methods.[146] For hydrothermal synthesis, the precursor synthesis, washing/

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

drying, lithiation, and calcination are considered for determining the input processing parameters of the procedure, such as reaction temperature, reaction time, precursor composition, impurity type and amount, washing time, drying temperature and time, pulverization pressure and time, precalcination temperature and time, and calcination temperature and time.

For solid-state synthesis, the precursor grinding, wet milling, solvent evaporation, pelletizing, and calcination steps are the three main processing steps. Here, the parameters of interest include the composition of the precursor mixture, grinding pressure, speed and time, size and number of zirconia balls, solvent amount, ball milling speed and time, evaporation temperature and time, size and volume of the pellet, pressure and time of the pressing machine for pelletization, calcination temperature, time and flow rate of injected gas, and partial pressure of oxygen.

The output structural parameters are the primary particle size, shape, secondary particle size, shape, crystal structure (crystal symmetry and lattice constants), and crystallinity (crystal *vs* amorphous) of the primary/secondary particles, valence states of the surface elements, and composition and its gradient in the primary and secondary particles.

Our example shows a simplified version of building up the processing−structure library. If we include the electrolyte, binder, separator, and anode altogether, the scope and size of the library will be much larger.

As such, generalizing the processing−structure library for any materials will be a very large job to complete. We envision that with the development of NLP techniques to read all the documents related to the processing of materials as well as image recognition techniques to extract processing parameters from tables and graphs in such documents, the task of generalizing the processing−structure library will be expedited. This library will be fed into an appropriate ML model to create a function we can use for the reverse engineering of the materials from their multiscale structures.

The spatial dimensions of the structure of materials is only part of what makes them tick. There is a spectrum of characteristic relaxation times associated with the various chemical and physical processes operating at the materials differing structural length scales, which adds the dimension of time. The resulting dynamic spatiotemporal hierarchy implies that any material at any time has structural features that have not yet reached equilibrium. This is the reason that the structure of materials depends on how they are made and what conditions they endure during processing.[1]

Even at equilibrium inside repeating units of crystal lattices and grains exist structural defects as they lower the Gibb's free energy of the total system, leading to the realization that structural defects have reasons to exist and play vital roles at all length scales. Defects can make or break materials in many ways. It is the dopant ion in perfect silicon crystals that has changed society through Moore's law. Carbon impurities in pure iron metal gave birth to the steel backbone of the industrial revolution. Although "defect tolerance" remains a central tenet of modern materials science and engineering for its commercial and safety importance, "defect engineering" is ascendant in the minds of many materials engineers. This is because defects at various hierarchical levels are a principal opportunity for controlling material behavior.

Therefore, the output structural features include the size, shape, and symmetry parameters of multiscale repeating and randomly distributed units as well as sparsely distributed defects in multiple dimensions as shown in Table 2.

One may ask the following question: "How can we deconvolute the complex multiscale processes occurring at the same time and extract processing−structure relations at each scale?"

The studies of direct e-beam effects on materials can provide useful data for this task. A large number of e-beam-induced phenomena have been reported on the ordering of oxygen vacancies,[147] formation of vacancies, and extended defects in 2D materials,[148] beam-induced migration of single interstitials,[149] and formation of individual chemical bonds.[150] If we can develop a processing tool that uses an e-beam as the main stimulus for the motion of atoms, then we can develop processing−structure relations at the atomic scale.

At the 1 μm scale of a structure, the structural changes during quenching (rapid cooling), whereby crystal grains transform and subdivide into a hierarchy of low-temperature crystalline phases, can be matched with the processing temperature and time as well as the quenching rate. In addition to these conventional processes, electrodeposition can also control the crystallinity of materials from a dendritic structure to an amorphous phase by means of the applied current and bath conditions.

We can also build upon the hierarchy of models based on materials science, applied mechanics, and quantum physics. Thermodynamic codes such as THERMOCALC enable designers to simulate the structure of metal alloys at the 10 μm scale, where the chemical partitioning between the liquid and solid phases evolves during solidification processing.

Lastly, once we build the structure−property and processing−structure models based on the experimental and simulation data, we will be able to integrate them to make a seamless processing−structure−property model. By using this model, we envision that we can input the desired sets of properties for materials and obtain the optimum processing recipe to make them.

M3I3 CHALLENGES AND FUTURE PERSPECTIVE

The greatest challenge of M3I3 is the quality and quantity of experimental data that could be used to develop the processing−structure−property relationship-based model.

Regarding the quality of the data, we can think of the following example of imaging the structure of a three-dimensional polycrystalline sample, where grains are defined by a single crystal in which the crystallographic planes are aligned in one direction and the grain boundaries act as interfaces between those grains. One simple method to build this image is to combine the use of SEM, focused ion beam (FIB) lithography, and electron backscatter diffraction (EBSD) altogether. In this case, we lose part of the materials due to the Ga beam impinging upon the materials of interest and suffer from Ga-ion implantation. The first challenge can be solved with the imputation of data using ML guided by physical and chemical laws. For example, the boundary should be continuous and its slope should be continuous unless there are more than two boundaries such as a triple boundary junction in the missing area. The distortion of grains or grain boundaries by Ga-ion beam damage can be detected if again ML is well-developed to identify such a region and modify the shape of grain boundaries based on the ML model.

Indeed, one of the breakthroughs ML brings to materials imaging is the imputation of missing data or non-existing data between existing data. One can then enhance a low-resolution image to a high-resolution image using ML, and this has been

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

Table 2. Multiscale Structural Components in Terms of Major (Repeating Units) and Minor (Randomly or Sparsely Distributed Units) Components and Their Features

size	major components	minor components	features
atom/ion	nucleus, electrons	holes, excess electrons, excess neutrons	electron density/spin, Fermi level, vacuum level, valence charge, size, mass, orbital shape
lattice/chain	atom, ion, molecule, bonding	vacancy, interstitial, dislocation, impurity	lattice parameter (size, angle), ionicity of bonding, bond length/angle, coordination number (CN), chain length/angle/conformation, octahedra/tetrahedra twist/tilt/rotation, tolerance factor, dislocation density/shape/size/charge
grain/particle	crystal lattice, molecular chain, core/shell	grain boundary, twin, precipitate, dopant	grain size/shape, grain boundary charge/width/length, core/shell ratio, composition, composition gradient, charge trap density
bulk	grain, particle, binder	additive, void	composition, composition gradient, interface size/length/shape, porosity, pore size/shape

shown in many movies where the actor asks the agent to enhance the quality of an image captured by CCTV on the street. The other breakthrough of ML is identifying and classifying objects in the image, which provides the possibility of autonomous driving based on the images captured by all cameras on the car.

The basic concept of convolutional neural network (CNN) used in object detection (localization and classification) is to learn the features of data and simplify data representations for the purpose of finding patterns. As the model learns, it simplifies features at each layer (edges, angles, and so on) and attribute a combination of features to a specific output (object class, location, and boundary). This is how ML get information about the location and speed of pedestrians, cars, and buildings in the image. Using the same ML structure and applying it to the materials images, one can segment the materials image into different phases and pores with information about location, size, and shape.

The ability to quickly analyze large imaging data sets is vital to the widespread adoption of modern materials characterization tools and, thus, the development of materials. Image segmentation can be the most subjective and time-consuming step in the data analysis workflow. A promising approach to segmentation of large materials data sets is the use of CNNs. For example, Horwath *et al.*[151] proposed methods for optimizing performance of TEM image segmentation using CNN. However, a major challenge is to obtain the images and segmentations needed for CNN training, since this requires segmentations performed by humans. We plan to use SegNet-based CNN, which is trained on simple phase field simulations, to segment experimental materials science data.

Furthermore, we can use ML to extract structural descriptors, such as the space group, crystal structure, *d*-spacing, and zone axis. For descriptors, such as the space group, crystal structure, and zone axis, we use logistic regression to assign each output to different classes (classification), whereas, for *d*-spacing, we use linear or nonlinear regression to assign each output to an obtained numerical value.

For example, the machine learning-assisted analysis of electron microscopy has been developed to analyze structural information of materials and defects. Ziatdinov *et al.* proposed a deep learning-based "weakly supervised" approach that uses information on the coordinates of atomic species in scanning transmission electron microscopy (STEM) images in order to analyze the atomic structure and detect a variety of defects automatically.[152] Li *et al.* demonstrated and evaluated an automated approach to detect dislocation loop in STEM images using contemporary ML computer vision and image analysis techniques, which achieved performance similar to the human average across the same data set.[153]

ML-assisted analysis of electron microscopy has identified the crystal structure of unknown materials. Ziletti *et al.* constructed a classification model based on a convolutional neural network to automatically determine the crystal class according to crystal symmetry.[154] They calculated the diffraction images to train the model from 100,000 simulated crystal structures including heavily defective ones. The trained model successfully identified the crystal class in the presence of defects.[154] Aguiar *et al.* classified experimental high-resolution STEM data as 230 space groups.[155] Two-dimensional fast Fourier transformed diffraction patterns were azimuthally integrated and converted into a one-dimensional profile. The pattern and profile provide the structural classification details using the deep-learning model approach.[155] Kaufmann *et al.* performed the crystal structure

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

determination of 14 Bravais lattice with electron backscatter diffraction (EBSD) data by constructing ResNet50 and Xception models.[156] They achieved accuracy of 93.5% for ResNet50 and 91.2% for Xception and confirmed good performance rather than traditional Hough transform.

ML-assisted analysis of electron microscopy has also revealed nanoscale dynamics. Yao *et al.* analyzed liquid-phase transmission electron microscopy (TEM) images of various gold nanostructures using U-Net neural network-based image segmentation to extract physical and chemical meaning.[157] They generated simulated TEM images for model training and showed the U-Net model has better performance than the conventional thresholding method in blurred liquid-phase TEM condition.

However, because of the quantity and the way data analysis is conducted, the vast majority of high-quality experimental data are left unreported and a very small fraction of this data are published, which renders the unreported data inaccessible to the research community. To tackle this problem effectively and promote the full utilization of high-resolution structural and property imaging data, the wide adoption of ML should be implemented. The development of unsupervised image analysis tools using high-performance computing platforms has demonstrated the ability to allow the full analysis of atomic configurations in 2D, high-resolution imaging data in real time.[158−160]

How should we tackle the issue of a small quantity of high-quality data? Data mining from patents and data sheets publicly available on the Web sites of companies in addition to journal papers can vastly increase the amount of high-quality data. Data augmentation within the boundaries of physical and chemical laws and the simulation of such data within the constraints of science can also help tackle this problem. One example includes the enumerated classes of atomic defects extracted from the deep convolutional neural network−Gaussian mixture model analysis of STEM movies.[38] We can find another example from the work of Frey *et al.*,[161] where they developed an approach based on deep transfer learning, machine learning, and first-principles calculations to address the issue of a small number of high-quality data. They used physics-informed featurization to generate a minimal description of defect structures and identified optimal point defects in 2D materials.[161]

In process of executing the M3I3 initiative, we have built the data-driven experimental design based on traditional NCM cathode materials (nickel, cobalt, and manganese). With this establishment, our future direction is to expand this idea for achieving even higher discharge capacity, which can be realized *via* Li-rich cathode. However, one of the major challenges is the limitation of available data that describe the Li-rich cathode properties. To mitigate this problem, we propose two solutions: First, we should build an ML-guided data generator for data augmentation. To ensure the quality of generated data, we will include a quality control feature that restricts only the data above-defined threshold to be accepted in the data pool. Thus, the data with larger size can be used for machine learning-driven experimental design. Second, we will use an ML method based on "transfer learning".[162] Since the NCM cathode database shares the common feature to that of Li-rich cathode, one may consider repurposing the NCM trained model in assisting the Li-rich prediction. With the pretrained model and transfer learning, we expect to achieve outstanding prediction in Li-rich cathode even with the small data set.

Our expectation is based on some of the related prior works that used generative adversarial networks (GAN) and variational autoencoders (VAEs) to address the issues of insufficient data. For example, Court *et al.*[163] reported an autoencoder-based generative deep-representation learning pipeline for geometrically optimized 3D crystal structures, which predicts the values of formation energy, total energy, band gap, bulk modulus, shear modulus, Poisson ratio, refractive index, and dielectric constant.

Ma *et al.*[164] developed a transfer learning strategy, which combines real and simulated data and the augmentation of training data in a data mining process. For a specific task of grain instance image segmentation, they generated synthetic data by fusing the images obtained from simulating the physical mechanism of grain formation and the image style information in real images extracted by the generative adversarial networks. Because the time required to perform grain simulation and generate synthetic data is much smaller than that to acquire real experimental data, their strategy can exploit the strong prediction power of deep learning without the experimental burden of training data preparation.

Ohno[165] addressed the issue of small data size for training models for regression models by using a real-valued nonvolume preserving model (real-NVP) as the normalizing flow and the generator in the GAN-based training method. Using kernel ridge regression trained by generated data, generalization performance was measured by evaluating the models, where the GAN-based training approach was comparable to the state-of-art models. Furthermore, Ohno used variational autoencoders (VAEs) as generative models for data augmentation to address the issue of small data size for regression problems.[166] The multitask learning for VAEs improved the generalization performance of multivariable linear regression model trained with augmented data in seven benchmark data sets and ionic conductivity data set.

While the most common approach for machine learning has been to extract physically meaningful descriptors and use either linear machine learning algorithms or kernel tricks to impose nonlinearity, the boom in large structured data sets and low-cost embarrassingly parallel computing has given rise to deep learning techniques in materials science.[167] Deep learning or neural network-based machine learning assembles large networks of neurons containing a matrix multiplication, the addition of a bias, and a nonlinearity. The weights and biases are optimized on the basis of some objective function. These networks if large enough are overparametrized to the point that they become universal function approximators. While overparametrization left unchecked will result in overfitting, the optimization is constrained through the addition of damaging mechanisms in the form of regularizers. Design of deep learning models requires co-design of the model to the objective. Generally, this involves building specific architectures for supervised learning (classification or regression), unsupervised latent space extraction, generative models, control systems, and much more. Deep learning has recently been used to extract latent manifolds from high-dimensional spectroscopy,[168] discover phase transformations,[169] segmentation, and detection in microscopy images,[170] and controlled experimentation[171] and atomic manipulation.[150]

Despite these successes, there are many open areas for innovation by M3I3. In particular, we are working on designing regularization pipeline mechanisms including adding sparsity through L1-regularization[168] and imposing selective Kullback−Leibler divergence constraints,[172] and shaping latent manifolds

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

using normalizing flows.[173] Additionally, we are working on creating models that can learn underlying governing equations that describe observable phenomena.[174] One of the key practical challenges with deep learning is the relatively long inference latency. There is a strong directive to use machine learning to analyze high-velocity streaming data > 1 TB/s and make decisions with faster than human response times (~300 ms). Both of these applications require alternative computing paradigms where analysis happens without saving data to disk and without CPU management.

The necessary innovation requires the co-design of deep learning algorithms and AI hardware for specific applications. One promising approach involves compressing deep learning models using pruning and quantization such that they can be deployed on reconfigurable logic such as field-programmable gate arrays capable of conducting inference with ns-latencies.[175]

Large volumes of searchable structural images and property maps will allow the building of a database of multiscale configurations and their associated properties in materials, thereby revealing the hierarchical physics and chemistry underlying the multiscale multimodal structure−property relationships in materials.

We plan to build an M3I3 repository of searchable structural and property maps using FAIR—findable, accessible, interoperable, and reusable—principles to standardize best practices as well as streamline the training of early career researchers.[176] Our repository will include data of all formats and a whole class of meta-analyses and connect with existing repositories such as NOMAD, JARVIS, Materials Project, AFLOW, OQMD, NREL MatDB, and Materials Cloud. The Materials Research Platform is the vision of a future system, which comprises data and knowledge assets, automation of science, and integrative approaches for materials research.[177] Many of the issues and topics such as inclusion of "dark data" (*i.e.*, data that are considered a "negative" result and not publishable), maintenance of data quality through rapid user feedback and community review, and curation of standardized and benchmarked data sets have been well-documented and are identical in substance when compared with M3I3. In the wider context, M3I3 is not a standalone initiative but rather a parallel and interdependent initiative, which needs a close connection to the Materials Research Platform.

With advances in experimental imaging and the availability of well-resolved information and big data, along with significant advances in high-performance computing and a worldwide thrust toward a general, collaborative, integrative, and on-demand research platform, there is a clear confluence in the required capabilities of advancing the M3I3 initiative.

Once we succeed in using the inverse "property−structure−processing" solver to develop cathode, anode, electrolyte, and membrane materials for high energy density Li-ion batteries, we will expand our scope of materials to, *e.g.*, battery/fuel cell, aerospace, automobile, food, medicine, and cosmetic materials.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.1c00211.

(Figure S1) Photograph of KAIST 2031 Vision Day in 2018 with M3I3 logo; (list and Table S1) M3I3 related initiatives such as MGI, NOMAD, MARVEL, MI^{2}I, MGE project, KAIST M3I3, and Data-based Materials Research Innovation; (Table S2) data in main text for Li-rich electrodes; (Figure S2) general molecular PVDF structure; (Figure S3) unit cell configuration of α-, β-, and γ-phases; (Figure S4) multiscale structure−property PVDF relationship of as an example of structure−property; (Figure S5) processing−structure relationships in molecular materials (PDF)

(Video S1) Full 3D tomograms of Figure 16 (MP4)

AUTHOR INFORMATION

Corresponding Author

Seungbum Hong − *Department of Materials Science and Engineering and KAIST Institute for NanoCentury (KINC), Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-2667-1983; Email: seungbum@kaist.ac.kr

Authors

Chi Hao Liow − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-7207-4848

Jong Min Yuk − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-4677-7363

Hye Ryung Byon − *Department of Chemistry, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-3692-6713

Yongsoo Yang − *Department of Physics, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-8654-302X

EunAe Cho − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-2871-6903

Jiwon Yeom − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-5183-0253

Gun Park − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-2496-3656

Hyeonmuk Kang − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-0174-555X

Seunggu Kim − *Department of Chemistry, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-3914-1295

Yoonsu Shim − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-6457-4623

Moony Na − *Department of Chemistry, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-7172-9494

Chaehwa Jeong − *Department of Physics, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-9153-710X

Gyuseong Hwang − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and*

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

Engineering (KAIST), Daejeon 34141, Republic of Korea; orcid.org/0000-0002-6902-9683
Hongjun Kim − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-2689-9769
Hoon Kim − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-7449-7406
Seongmun Eom − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-2306-021X
Seongwoo Cho − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-3000-4251
Hosun Jun − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-1863-1391
Yongju Lee − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-8737-2477
Arthur Baucour − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-8251-5504
Kihoon Bang − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-5067-034X
Myungjoon Kim − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-1331-8854
Seokjung Yun − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-9992-3395
Jeongjae Ryu − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-8052-9059
Youngjoon Han − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-6180-6022
Albina Jetybayeva − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-0652-3756
Pyuck-Pa Choi − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0001-9920-0755
Joshua C. Agar − *Department of Materials Science and Engineering, Lehigh University, Bethlehem, Pennsylvania 18015, United States;* orcid.org/0000-0001-5411-4693
Sergei V. Kalinin − *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States;* orcid.org/0000-0001-5354-6152
Peter W. Voorhees − *Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, United States*
Peter Littlewood − *James Franck Institute, University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0003-0282-9325
Hyuck Mo Lee − *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Engineering (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-4556-6692

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.1c00211

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the following members for their input in gathering the data sheet and critical comments on the manuscript: Aditi Saha, Jaewoon Kim, Soyeon Kim, Eunnuri Cho, Hyunji Kim, Youngwoo Choi, Chungik Oh, Jaegyu Kim, Jimin Oh, Hyoungkyu Kim, Jaewook Shin, DaeHee Lee, Ho Jun Lee, Jae Yeol Park, Han Beom Jeong, Jae Sang Lee, Joon Ha Chang, Yohan Kim, Sujung Kim, Hyunjeong Oh, Jae-Wook Han, Kyuseon Jang, Boryung Yoo, Hyeonjin Park, Minhwan Cho, Jun Hyung Park, Seokhwan Min, Hyesung Jo, Yeeun Kim, Jung Woo Choi, Youngtae Park, and Doosun Hong. We thank our internal and external advisory board members (Prof. Jonghwa Shin (MSE), Prof. Se-Young Yun (AI), Prof. Chang Dong Yoo (EE), Prof. Hye Won Chung (EE), and Prof. Myung Joon Han (Physics) at KAIST and Dr. Roger Proksch at Asylum Research) for their advice and critical comments on the M3I3 project. This work was supported by the KAIST-funded Global Singularity Research Program for 2019 and 2020. J.C.A. acknowledges support from the National Science Foundation under Grant TRIPODS + X:RES-1839234 and the Nano/Human Interfaces Presidential Initiative. S.V.K.'s effort was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Materials Sciences and Engineering Division and was performed at the Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS), a U.S. Department of Energy, Office of Science User Facility.

VOCABULARY

Materials by Design, a methodology to design a material that meets a material user's need; Materials Genome Initiative, an initiative to design, manufacture, and deploy materials and materials-based technologies faster and cheaper than ever before; materials informatics, a field of study that applies the principles of informatics to materials science and engineering to improve the understanding, design, development, and discovery of materials; multimodal imaging, imaging materials in various modes that can map structures and properties using different descriptors or markers (for example, piezoresponse force microscopy uses piezoelectric strain as the marker for polarization vector in ferroelectric materials); multiscale imaging, imaging materials at various scales covering atomic scale, nanoscale, mesoscale, micrometer scale, millimeter scale, and meter scale; ten flagship research fields, grand research fields to

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

overcome the global challenges, which can be categorized into four major themes, namely, bio/medical field, energy and environment, military/defense/space, and the fourth industrial revolution

REFERENCES

(1) Olson, G. B. Designing a New Material World. *Science* **2000**, *288*, 993−998.
(2) Xiong, W.; Olson, G. B. Cybermaterials: Materials by Design and Accelerated Insertion of Materials. *npj. Comput. Mater.* **2016**, *2*, 15009.
(3) *Materials Genome Initiative*; http://mgi.gov (accessed 2020-10-18).
(4) *"Material Research by Information Integration" Initiative (MI²I)*; https://www.nims.go.jp/MII-I/en/ (accessed 2020-10-18).
(5) *Horizon 2020: Novel Materials Discovery*; https://cordis.europa.eu/project/id/951786 (accessed 2020-10-18).
(6) *Vom Material zur Innovation (From Material to Innovation)*. Bundesministerium für Bildung und Forshung; https://www.bmbf.de/de/vom-material-zur-innovation-1130.html (accessed 2020-10-18).
(7) *Materials Scientific Data Sharing Network*, http://www.materdata.cn/ (accessed 2020-11-04).
(8) *Creative Materials Discovery Program*; https://nrf.re.kr/biz/info/notice/list?biz_no=305 (accessed 2020-10-18).
(9) Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The Throughput Highway to Computational Materials Design. *Nat. Mater.* **2013**, *12*, 191−201.
(10) Lopez-Bezanilla, A.; Littlewood, P. B. Growing Field of Materials Informatics: Databases and Artificial Intelligence. *MRS Commun.* **2020**, *10*, 1−10.
(11) Kalinin, S. V.; Sumpter, B. G.; Archibald, R. K. Big-Deep-Smart Data in Imaging for Guiding Materials Design. *Nat. Mater.* **2015**, *14*, 973−980.
(12) KAIST Vision 2031 Committee; Gimmyoung, P. *2031 KAIST Future Report* (in Korean); 2018; pp 1−276, http://www.gimmyoung.com/Book/BookView?bookCode=BC002877&bookType=all (accessed 2020-10-18).
(13) Under the Heading of Global Singularity Research Projects, M3I3 Initiative can be found: M313 Initiative: Materialsand Molecular Modeling, Imaging, Informatics and Integration. *KAIST Research Homepage*, KAIST, Daejon, Republic of Korea; https://www.kaist.ac.kr/en/html/research/04.html (accessed 2020-10-18).
(14) Opening New Horizons for Humanity: KAIST. *Nature Advertorial*, Nature Research Custom Media, 2020; https://www.nature.com/articles/d42473-020-00132-w (accessed 2020-10-18).
(15) *Materials Science*, Wikipedia. https://en.wikipedia.org/wiki/Materials_science (accessed 202010-18).
(16) Materials Age. *SciMAP*, Mar. 20, 2018; https://www.scimap.tech/post/materials-age (accessed 2021-01-14).
(17) Gibbs, J. W. On the Equilibrium of Heterogeneous Substances. *Transactions of the Connecticut Academy of Arts and Sciences*; Connecticut Academy of Arts and Sciences, 1874−1878; Vol. *3*, pp 108−248 and 343−524.
(18) Mody, C. C. M. *The Long Arm of Moore's Law Microelectronics and American Science*. The MIT Press: Cambridge, MA, USA, 2017; pp 1−304.
(19) Riordan, M.; Hoddeson, L. *Crystal Fire: The Birth of the Information Age*; W. W. Norton: New York, NY, USA, 1997; pp 1−352.
(20) Smith, C. S. *A History of Metallography*; The MIT Press: Cambridge, MA, USA, 1988; pp 1−326.
(21) Hall, E. O. The Deformation and Ageing of Mild Steel: III Discussion of Results. *Proc. Phys. Soc., London, Sect. B* **1951**, *64*, 747−753.
(22) *The Nobel Prize in Physics 1986*;https://www.nobelprize.org/prizes/physics/1986/summary/ (accessed 2021-01-21).
(23) Williams, D. B.; Carter, C. B., The Transmission Electron Microscope. *Transmission Electron Microscopy: A Textbook for Materials Science*; Springer: Boston, MA, USA, 1996; pp 3−17, DOI: 10.1007/978-0-387-76501-3_1.
(24) Palucka, T. Overview of Electron Microscopy; https://authors.library.caltech.edu/5456/1/hrst.mit.edu/hrs/materials/public/ElectronMicroscope/EM_HistOverview.htm (accessed 2021-01-21).
(25) Kim, H.; Lee, S.; Kim, S.; Oh, C.; Ryu, J.; Kim, J.; Park, E.; Hong, S.; No, K. Membrane Crystallinity and Fuel Crossover in Direct Ethanol Fuel Cells with Nafion Composite Membranes Containing Phosphotungstic Acid. *J. Mater. Sci.* **2017**, *52*, 2400−2412.
(26) About Jung Ho Kim. *Encyclopedia of Korean Culture*, The Academy of Korean Studies; http://encykorea.aks.ac.kr/Contents/Index?contents_id=E0010423 (accessed 2020-10-18).
(27) Daedongyeo Map. Full image. *Wikimedia*; https://commons.wikimedia.org/wiki/File:Daedongyeojido-full.jpg (accessed 202010-19).
(28) Choi, Y.-Y.; Sharma, P.; Phatak, C.; Gosztola, D. J.; Liu, Y.; Lee, J.; Lee, B.; Li, J.; Gruverman, A.; Ducharme, S.; Hong, S. Enhancement of Local Piezoresponse in Polymer Ferroelectrics via Nanoscale Control of Microstructure. *ACS Nano* **2015**, *9*, 1809−1819.
(29) Mount Baekdu. *Google Maps*; https://bit.ly/344HVt0 (accessed 2020-10-18).
(30) Mount Everest. *Google Maps*; https://bit.ly/2HbL8Ok (accessed 2020-10-18).
(31) Wang, L. H. Picture of Mount Everest. This file is licensed under the Creative Commons Attribution-Share Alike 4.0 International licensehttps://commons.wikimedia.org/wiki/File:Mt.Everest.jpg (accessed 2020-10-19).
(32) Balke, N.; Jesse, S.; Morozovska, A. N.; Eliseev, E.; Chung, D. W.; Kim, Y.; Adamczyk, L.; Garcia, R. E.; Dudney, N.; Kalinin, S. V. Nanoscale Mapping of ion Diffusion in a Lithium-Ion Battery Cathode. *Nat. Nanotechnol.* **2010**, *5*, 749−754.
(33) Balke, N.; Kalnaus, S.; Dudney, N. J.; Daniel, C.; Jesse, S.; Kalinin, S. V. Local Detection of Activation Energy for Ionic Transport in Lithium Cobalt Oxide. *Nano Lett.* **2012**, *12*, 3399−3403.
(34) Yang, S.; Yan, B.; Li, T.; Zhu, J.; Lu, L.; Zeng, K. In Situ Studies of Lithium-Ion Diffusion in a Lithium-Rich Thin Film Cathode by Scanning Probe Microscopy Techniques. *Phys. Chem. Chem. Phys.* **2015**, *17*, 22235−22242.
(35) Alikin, D. O.; Romanyuk, K. N.; Slautin, B. N.; Rosato, D.; Shur, V. Y.; Kholkin, A. L. Quantitative Characterization of the Ionic Mobility and Concentration in Li-Battery Cathodes via Low Frequency Electrochemical Strain Microscopy. *Nanoscale* **2018**, *10*, 2503−2511.
(36) Kim, H.; Oh, J.; Park, G.; Jetybayeva, A.; Kim, J.; Lee, Y.-G.; Hong, S. Visualization of Functional Components in a Lithium Silicon Titanium Phosphate-Natural Graphite Composite Anode. *ACS Appl. Energy Mater.* **2020**, *3*, 3253−3261.
(37) Kim, H.; Yun, S.; Kim, K.; Kim, W.; Ryu, J.; Nam, H. G.; Han, S. M.; Jeon, S.; Hong, S. Breaking the Elastic Limit of Piezoelectric Ceramics using Nanostructures: A Case Study using ZnO. *Nano Energy* **2020**, *78*, 105259.
(38) Kalinin, S. V.; Dyck, O.; Balke, N.; Neumayer, S.; Tsai, W.-Y.; Vasudevan, R.; Lingerfelt, D.; Ahmadi, M.; Ziatdinov, M.; McDowell, M. T.; Strelcov, E. Toward Electrochemical Studies on the Nanometer and Atomic Scales: Progress, Challenges, and Opportunities. *ACS Nano* **2019**, *13*, 9735−9780.
(39) Hong, S.; Nakhmanson, S. M.; Fong, D. D. Screening Mechanisms at Polar Oxide Heterointerfaces. *Rep. Prog. Phys.* **2016**, *79*, 076501.
(40) Hong, S., Ed. *Nanoscale Phenomena in Ferroelectric Thin Films*; Kluwer Academic: Boston, MA, USA, 2004; pp 3−279.
(41) Auciello, O.; Scott, J. F.; Ramesh, R. The Physics of Ferroelectric Memories. *Phys. Today* **1998**, *51*, 22−27.
(42) Setter, N.; Damjanovic, D.; Eng, L.; Fox, G.; Gevorgian, S.; Hong, S.; Kingon, A.; Kohlstedt, H.; Park, N. Y.; Stephenson, G. B.; Stolitchnov, I.; Taganstev, A. K.; Taylor, D. V.; Yamada, T.; Streiffer, S. Ferroelectric Thin Films: Review of Materials, Properties, and Applications. *J. Appl. Phys.* **2006**, *100*, 051606.
(43) Auciello, O.; Araujo, C. A. P. d.; Celinska, J. Review of the Science and Technology for Low- and High-Density Nonvolatile Ferroelectric Memories. *Emerging Non-Volatile Memories*: Hong, S., Auciello, O.,

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

Wouters, D., Eds.; Springer: New York, NY, USA, 2014; pp 3−36, DOI: 10.1007/978-1-4899-7537-9_1.
(44) Fong, D. D.; Stephenson, G. B.; Streiffer, S. K.; Eastman, J. A.; Auciello, O.; Fuoss, P. H.; Thompson, C. Ferroelectricity in Ultrathin Perovskite Films. *Science* **2004**, *304*, 1650−1653.
(45) Tanaka, K.; Kurihashi, Y.; Uda, T.; Daimon, Y.; Odagawa, N.; Hirose, R.; Hiranaga, Y.; Cho, Y. Scanning Nonlinear Dielectric Microscopy Nano-Science and Technology for Next Generation High Density Ferroelectric Data Storage. *Jpn. J. Appl. Phys.* **2008**, *47*, 3311−3325.
(46) Cho, Y.; Hong, S. Scanning Probe-Type Data Storage Beyond Hard Disk Drive and Flash Memory. *MRS Bull.* **2018**, *43*, 365−369.
(47) Hong, S.; Colla, E. L.; Kim, E.; Taylor, D. V.; Tagantsev, A. K.; Muralt, P.; No, K.; Setter, N. High Resolution Study of Domain Nucleation and Growth During Polarization Switching in Pb(Zr, Ti)O3 Ferroelectric Thin Film Capacitors. *J. Appl. Phys.* **1999**, *86*, 607−613.
(48) Hong, S.; Setter, N. Evidence for Forward Domain Growth being Rate-Limiting Step in Polarization Switching in < 111>-Oriented-Pb(Zr0.45Ti0.55)O3 Thin-Film Capacitors. *Appl. Phys. Lett.* **2002**, *81*, 3437−3439.
(49) Colla, E. L.; Hong, S.; Taylor, D. V.; Tagantsev, A. K.; Setter, N.; No, K. Direct Observation of Region by Region Suppression of the Switchable Polarization (Fatigue) in Pb(Zr, Ti)O3 Thin Film Capacitors with Pt Electrodes. *Appl. Phys. Lett.* **1998**, *72*, 2763−2765.
(50) Takata, K. Strain Imaging of a Pb(Zr,Ti)O3 Thin Film. *J. Appl. Phys.* **1996**, *79* (1), 134−142.
(51) Kolosov, O.; Gruverman, A.; Hatano, J.; Takahashi, K.; Tokumoto, H. Nanoscale Visualization and Control of Ferroelectric Domains by Atomic Force Microscopy. *Phys. Rev. Lett.* **1995**, *74*, 4309−4312.
(52) Franke, K.; Hülz, H.; Weihnacht, M.; Häbler, W.; Besold, J. Nanoscale Investigations of Polarization in Thin Ferroelectric Films by Means of Scanning Force Microscopy. *Ferroelectrics* **1995**, *172*, 397−404.
(53) Franke, K.; Besold, J.; Haessler, W.; Seegebarth, C. Modification and Detection of Domains on Ferroelectric PZT Films by Scanning Force Microscopy. *Surf. Sci.* **1994**, *302*, L283−L288.
(54) Hong, J. W.; Park, S.-I.; Khim, Z. G. Measurement of Hardness, Surface Potential, and Charge Distribution with Dynamic Contact Mode Electrostatic Force Microscope. *Rev. Sci. Instrum.* **1999**, *70*, 1735−1739.
(55) Balke, N.; Bdikin, I.; Kalinin, S. V.; Kholkin, A. L. Electromechanical Imaging and Spectroscopy of Ferroelectric and Piezoelectric Materials: State of the Art and Prospects for the Future. *J. Am. Ceram. Soc.* **2009**, *92*, 1629−1647.
(56) Hong, S.; Woo, J.; Shin, H.; Jeon, J. U.; Pak, Y. E.; Colla, E. L.; Setter, N.; Kim, E.; No, K. Principle of Ferroelectric Domain Imaging using Atomic Force Microscope. *J. Appl. Phys.* **2001**, *89*, 1377−1386.
(57) Park, H.; Jung, J.; Min, D. K.; Kim, S.; Hong, S.; Shin, H. Scanning Resistive Probe Microscopy: Imaging Ferroelectric Domains. *Appl. Phys. Lett.* **2004**, *84*, 1734−1736.
(58) Ko, H.; Ryu, K.; Park, H.; Park, C.; Jeon, D.; Kim, Y. K.; Jung, J.; Min, D. K.; Kim, Y.; Lee, H. N.; Park, Y.; Shin, H.; Hong, S. High-Resolution Field Effect Sensing of Ferroelectric Charges. *Nano Lett.* **2011**, *11*, 1428−1433.
(59) Hong, S.; Tong, S.; Park, W. I.; Hiranaga, Y.; Cho, Y.; Roelofs, A. Charge Gradient Microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 6566−6569.
(60) Tong, S.; Jung, I. W.; Choi, Y. Y.; Hong, S.; Roelofs, A. Imaging Ferroelectric Domains and Domain Walls using Charge Gradient Microscopy: Role of Screening Charges. *ACS Nano* **2016**, *10*, 2568−2574.
(61) Hong, S.; Kim, Y. Ferroelectric Probe Storage Devices. *Emerging Non-Volatile Memories*: Hong, S., Auciello, O., Wouters, D., Eds.; Springer: New York, NY, USA, 2014; 259−273, DOI: 10.1007/978-1-4899-7537-9_7.
(62) Hong, S.; Choa, S.; Jung, J.; Ko, H.; Kim, Y. K. Ferroelectric Hard Disk System. U.S. Pat. 8248906, 2012.
(63) Integrated Imaging Institute (I3). *Argonne National Laboratory*; https://www.anl.gov/imaging (accessed 2020-10-18).
(64) Institute for Functional Imaging of Materials (IFIM). *Oak Ridge National Laboratory*; https://www.ornl.gov/facility/cnms (accessed 2020-11-04).
(65) Cheng, L.; Assary, R. S.; Qu, X.; Jain, A.; Ong, S. P.; Rajput, N. N.; Persson, K.; Curtiss, L. A. Accelerating Electrolyte Discovery for Energy Storage with High-Throughput Screening. *J. Phys. Chem. Lett.* **2015**, *6*, 283−291.
(66) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.
(67) de Pablo, J. J.; Jackson, N. E.; Webb, M. A.; Chen, L.-Q.; Moore, J. E.; Morgan, D.; Jacobs, R.; Pollock, T.; Schlom, D. G.; Toberer, E. S.; Analytis, J.; Dabo, I.; DeLongchamp, D. M.; Fiete, G. A.; Grason, G. M.; Hautier, G.; Mo, Y.; Rajan, K.; Reed, E. J.; Rodriguez, E.; Stevanovic, V.; Suntivich, J.; Thornton, K.; Zhao, J.-C. New Frontiers for the Materials Genome Initiative. *npj Comput. Mater.* **2019**, *5*, 41.
(68) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P. A Critical Review of Machine Learning of Energy Materials. *Adv. Energy Mater.* **2020**, *10*, 1903242.
(69) Ng, A. Machine Learning. *Coursera*; https://www.coursera.org/learn/machine-learning (accessed 2020-10-18).
(70) Legendre, A. M. *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*; Courcier: Paris, 1805; pp 1−98.
(71) Samuel, A. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* **2000**, *44*, 206−226.
(72) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533−536.
(73) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436−444.
(74) Chibani, S.; Coudert, F.-X. Machine Learning Approaches for the Prediction of Materials Properties. *APL Mater.* **2020**, *8*, 080701.
(75) Attia, P. M.; Grover, A.; Jin, N.; Severson, K. A.; Markov, T. M.; Liao, Y. H.; Chen, M. H.; Cheong, B.; Perkins, N.; Yang, Z.; Herring, P. K.; Aykol, M.; Harris, S. J.; Braatz, R. D.; Ermon, S.; Chueh, W. C. Closed-Loop Optimization of Fast-Charging Protocols for Batteries with Machine Learning. *Nature* **2020**, *578*, 397−402.
(76) Min, K.; Choi, B.; Park, K.; Cho, E. Machine Learning Assisted Optimization of Electrochemical Properties for Ni-Rich Cathode Materials. *Sci. Rep.* **2018**, *8*, 15778.
(77) Raccuglia, P.; Elbert, K. C.; Adler, P. D. F.; Falk, C.; Wenny, M. B.; Mollo, A.; Zeller, M.; Friedler, S. A.; Schrier, J.; Norquist, A. J. Machine-Learning-Assisted Materials Discovery using Failed Experiments. *Nature* **2016**, *533*, 73−76.
(78) Zakutayev, A.; Wunder, N.; Schwarting, M.; Perkins, J. D.; White, R.; Munch, K.; Tumas, W.; Phillips, C. An Open Experimental Database for Exploring Inorganic Materials. *Sci. Data* **2018**, *5*, 180053.
(79) Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S.; Grewe, D.; Nham, J.; Kalchbrenner, N.; Sutskever, I.; Lillicrap, T.; Leach, M.; Kavukcuoglu, K.; Graepel, T.; Hassabis, D. Mastering The Game of Go with Deep Neural Networks and Tree Search. *Nature* **2016**, *529*, 484−489.
(80) Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A.; Chen, Y.; Lillicrap, T.; Hui, F.; Sifre, L.; van den Driessche, G.; Graepel, T.; Hassabis, D. Mastering the Game of Go without Human Knowledge. *Nature* **2017**, *550*, 354−359.
(81) Noh, J.; Gu, G. H.; Kim, S.; Jung, Y. Machine-Enabled Inverse Design of Inorganic Solid Materials: Promises and Challenges. *Chem. Sci.* **2020**, *11*, 4871−4881.
(82) *The Materials Project*; https://materialsproject.org/ (accessed 2020-10-18).
(83) *The Citrination Platform*; https://citrination.com/ (accessed 2020-11-04).

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

(84) O'Mara, J.; Meredig, B.; Michel, K. Materials Data Infrastructure: A Case Study of the Citrination Platform to Examine Data Import, Storage, and Access. *JOM* **2016**, *68*, 2031−2034.
(85) Huber, S. P.; Zoupanos, S.; Uhrin, M.; Talirz, L.; Kahle, L.; Häuselmann, R.; Gresch, D.; Müller, T.; Yakutovich, A. V.; Andersen, C. W.; Ramirez, F. F.; Adorf, C. S.; Gargiulo, F.; Kumbhar, S.; Passaro, E.; Johnston, C.; Merkys, A.; Cepellotti, A.; Mounet, N.; Marzari, N.; Kozinsky, B.; Pizzi, G.; et al. AiiDA 1.0, A Scalable Computational Infrastructure for Automated Reproducible Workflows and Data Provenance. *Sci. Data* **2020**, *7*, 300.
(86) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Comput. Mater. Sci.* **2012**, *58*, 218−226.
(87) *AFLOW (Automatic-FLOW for Materials Discovery)*; http://www.aflow.org/ (accessed 2020-10-18).
(88) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Comput. Mater.* **2015**, *1*, 15010.
(89) *The Open Quantum Materials Database (OQMD)*; See http://oqmd.org/ (accessed 2020-10-18).
(90) *NOMAD Repository & Archive*; https://nomad-repository.eu/ (accessed 2020-10-18).
(91) Choudhary, K.; Zhang, Q.; Reid, A. C. E.; Chowdhury, S.; Van Nguyen, N.; Trautt, Z.; Newrock, M. W.; Congo, F. Y.; Tavazza, F. Computational Screening of High-Performance Optoelectronic Materials Using Optb88vdw and TB-mBJ Formalisms. *Sci. Data* **2018**, *5*, 180082.
(92) Choudhary, K.; Garrity, K. F.; Reid, A. C. E.; DeCost, B.; Biacchi, A. J.; Hight Walker, A. R.; Trautt, Z.; Hattrick-Simpers, J.; Kusne, A. G.; Centrone, A.; Davydov, A.; Jiang, J.; Pachter, R.; Cheon, G.; Reed, E.; Agrawal, A.; Qian, X.; Sharma, V.; Zhuang, H.; Kalinin, S. V.; Sumpter, B. G.; Pilania, G.; Acar, P.; Mandal, S.; Haule, K.; Vanderbilt, D.; Rabe, K.; Tavazza, F.; et al. The Joint Automated Repository for Various Integrated Simulations (JARVIS) for Data-Driven Materials Design. *npj Comput. Mater.* **2020**, *6*, 173.
(93) Belsky, A.; Hellenbrandt, M.; Karen, V. L.; Luksch, P. New Developments in the Inorganic Crystal Structure Database (ICSD): Accessibility in Support of Materials Research and Design. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 364−369.
(94) Blaiszik, B.; Chard, K.; Pruyne, J.; Ananthakrishnan, R.; Tuecke, S.; Foster, I. The Materials Data Facility: Data Services to Advance Materials Science Research. *JOM* **2016**, *68*, 2045−2052.
(95) Alberi, K.; Nardelli, M. B.; Zakutayev, A.; Mitas, L.; Curtarolo, S.; Jain, A.; Fornari, M.; Marzari, N.; Takeuchi, I.; Green, M. L.; Kanatzidis, M.; Toney, M. F.; Butenko, S.; Meredig, B.; Lany, S.; Kattner, U.; Davydov, A.; Toberer, E. S.; Stevanovic, V.; Walsh, A.; Park, N.-G.; Aspuru-Guzik, A.; Tabor, D. P.; Nelson, J.; Murphy, J.; Setlur, A.; Gregoire, J.; Li, H.; Xiao, R.; Ludwig, A.; Martin, L. W.; Rappe, A. M.; Wei, S.-H.; Perkins, J.; et al. The 2019 Materials by Design Roadmap. *J. Phys. D: Appl. Phys.* **2019**, *52*, 013001.
(96) Mathew, K.; Montoya, J. H.; Faghaninia, A.; Dwarakanath, S.; Aykol, M.; Tang, H.; Chu, I.-H.; Smidt, T.; Bocklund, B.; Horton, M.; Dagdelen, J.; Wood, B.; Liu, Z.-K.; Neaton, J.; Ong, S. P.; Persson, K.; Jain, A. Atomate: A High-Level Interface to Generate, Execute, and Analyze Computational Materials Science Workflows. *Comput. Mater. Sci.* **2017**, *139*, 140−152.
(97) Phatak, C.; Gaft, J.; Foster, I.; Blaiszik, B.Automated Data Curation for Electron Microscopy using the Materials Data Facility, https://www.microscopy.org/MandM/2019/program/abstracts/PDP-33.pdf (accessed 2020-12-26).
(98) Saal, J. E.; Oliynyk, A. O.; Meredig, B. Machine Learning in Materials Discovery: Confirmed Predictions and Their Underlying Approaches. *Annu. Rev. Mater. Res.* **2020**, *50*, 49−69.
(99) Chang, S. H.; Kim, J.; Phatak, C.; D'Aquila, K.; Kim, S. K.; Kim, J.; Song, S. J.; Hwang, C. S.; Eastman, J. A.; Freeland, J. W.; Hong, S. X-ray Irradiation Induced Reversible Resistance Change in Pt/TiO2/Pt Cells. *ACS Nano* **2014**, *8*, 1584−1589.
(100) Ziatdinov, M.; Dyck, O.; Sumpter, B. G.; Jesse, S.; Vasudevan, R. K.; Kalinin, S. V., Building and Exploring Libraries of Atomic Defects in Graphene: Scanning Transmission Electron and Scanning Tunneling Microscopy Study. *arXiv (Materials Science)*, 2019, arXiv:1809.04256. https://arxiv.org/abs/1809.04256 (accessed 2020-10-18).
(101) Goodenough, J. B.; Park, K.-S. The Li-Ion Rechargeable Battery: A Perspective. *J. Am. Chem. Soc.* **2013**, *135*, 1167−1176.
(102) Li, M.; Lu, J.; Chen, Z.; Amine, K. 30 Years of Lithium-Ion Batteries. *Adv. Mater.* **2018**, *30*, 1800561.
(103) Manthiram, A. A Reflection on Lithium-Ion Battery Cathode Chemistry. *Nat. Commun.* **2020**, *11*, 1550.
(104) Lithium-Ion Battery Market Size and Forecast. *Verified Market Research*, August 2020; https://www.verifiedmarketresearch.com/product/lithium-ion-battery-market/ (accessed 2020-10-18).
(105) Ji, X.; Lee, K. T.; Nazar, L. F. A Highly Ordered Nanostructured Carbon-Sulphur Cathode for Lithium-Sulphur Batteries. *Nat. Mater.* **2009**, *8*, 500−506.
(106) Dutta, A.; Wong, R. A.; Park, W.; Yamanaka, K.; Ohta, T.; Jung, Y.; Byon, H. R. Nanostructuring One-Dimensional and Amorphous Lithium Peroxide for High Round-Trip Efficiency in Lithium-Oxygen Batteries. *Nat. Commun.* **2018**, *9*, 680.
(107) Park, S.-K.; Lee, J.; Hwang, T.; Jang, B.; Piao, Y. Scalable Synthesis of Honeycomb-like Ordered Mesoporous Carbon Nanosheets and Their Application in Lithium-Sulfur Batteries. *ACS Appl. Mater. Interfaces* **2017**, *9*, 2430−2438.
(108) Zheng, J.; Guo, G.; Li, H.; Wang, L.; Wang, B.; Yu, H.; Yan, Y.; Yang, D.; Dong, A. Elaborately Designed Micro-Mesoporous Graphitic Carbon Spheres as Efficient Polysulfide Reservoir for Lithium-Sulfur Batteries. *ACS Energy Lett.* **2017**, *2*, 1105−1114.
(109) Feynman, R. P.; Leighton, R. B.; Sands, M. The Feynman Lectures in Physics 1963, II. *Mainly Electromagnetism and Matter*; Addison-Wesley: Reading, MA, USA, 1963; Chapters 1−29.
(110) Kononova, O.; Huo, H.; He, T.; Rong, Z.; Botari, T.; Sun, W.; Tshitoyan, V.; Ceder, G. Text-Mined Dataset of Inorganic Materials Synthesis Recipes. *Sci. Data* **2019**, *6*, 203.
(111) Composition Energy Density Relation. *M3I3 Google Data Sheet*; https://bit.ly/3nZrO7U (accessed 2020-10-19).
(112) Kumar Nayak, P.; Grinblat, J.; Levi, E.; Penki, T. R.; Levi, M.; Sun, Y.-K.; Markovsky, B.; Aurbach, D. Remarkably Improved Electrochemical Performance of Li- and Mn-Rich Cathodes upon Substitution of Mn with Ni. *ACS Appl. Mater. Interfaces* **2017**, *9*, 4309−4319.
(113) Kalinin, S. V.; Strelcov, E.; Belianinov, A.; Somnath, S.; Vasudevan, R. K.; Lingerfelt, E. J.; Archibald, R. K.; Chen, C.; Proksch, R.; Laanait, N.; Jesse, S. Big, Deep, and Smart Data in Scanning Probe Microscopy. *ACS Nano* **2016**, *10*, 9068−9086.
(114) Collins, L.; Liu, Y.; Ovchinnikova, O. S.; Proksch, R. Quantitative Electromechanical Atomic Force Microscopy. *ACS Nano* **2019**, *13*, 8055−8066.
(115) Yuk, J. M.; Zhou, Q.; Chang, J.; Ercius, P.; Alivisatos, A. P.; Zettl, A. Real-Time Observation of Water-Soluble Mineral Precipitation in Aqueous Solution by in Situ High-Resolution Electron Microscopy. *ACS Nano* **2016**, *10*, 88−92.
(116) Xu, Z.-L.; Kim, S. J.; Chang, D.; Park, K.-Y.; Dae, K. S.; Dao, K. P.; Yuk, J. M.; Kang, K. Visualization of Regulated Nucleation and Growth of Lithium Sulfides for High Energy Lithium Sulfur Batteries. *Energy Environ. Sci.* **2019**, *12*, 3144−3155.
(117) Song, K.; Ryu, S.; Lee, H.; Paudel, T. R.; Koch, C. T.; Park, B.; Lee, J. K.; Choi, S.-Y.; Kim, Y.-M.; Kim, J. C.; Jeong, H. Y.; Rzchowski, M. S.; Tsymbal, E. Y.; Eom, C.-B.; Oh, S. H. Direct Imaging of the Electron Liquid at Oxide Interfaces. *Nat. Nanotechnol.* **2018**, *13*, 198−203.
(118) Yang, Y.; Chen, C.-C.; Scott, M. C.; Ophus, C.; Xu, R.; Pryor, A.; Wu, L.; Sun, F.; Theis, W.; Zhou, J.; Eisenbach, M.; Kent, P. R. C.; Sabirianov, R. F.; Zeng, H.; Ercius, P.; Miao, J. Deciphering Chemical Order/Disorder and Material Properties at the Single-Atom Level. *Nature* **2017**, *542*, 75−79.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995

(119) Li, Z.; Li, Y.; Zhao, Y.; Wang, H.; Zhang, Y.; Song, B.; Li, X.; Lu, S.; Hao, X.-Q.; Hla, S.-W.; Tu, Y.; Li, X. Synthesis of Metallopolymers and Direct Visualization of the Single Polymer Chain. *J. Am. Chem. Soc.* **2020**, *142*, 6196−6205.
(120) Latt, K. Z.; Schlueter, J. A.; Darancet, P.; Hla, S.-W. Two-Dimensional Molecular Charge Density Waves in Single-Layer-Thick Islands of a Dirac Fermion System. *ACS Nano* **2020**, *14*, 8887−8893.
(121) Yao, M. J.; Welsch, E.; Ponge, D.; Haghighat, S. M. H.; Sandlöbes, S.; Choi, P.; Herbig, M.; Bleskov, I.; Hickel, T.; Lipinska-Chwalek, M.; Shanthraj, P.; Scheu, C.; Zaefferer, S.; Gault, B.; Raabe, D. Strengthening and Strain Hardening Mechanisms in a Precipitation-Hardened High-Mn Lightweight Steel. *Acta Mater.* **2017**, *140*, 258−273.
(122) Lin, R.; Hu, E.; Liu, M.; Wang, Y.; Cheng, H.; Wu, J.; Zheng, J.-C.; Wu, Q.; Bak, S.; Tong, X.; Zhang, R.; Yang, W.; Persson, K. A.; Yu, X.; Yang, X.-Q.; Xin, H. L. Anomalous Metal Segregation in Lithium-Rich Material Provides Design Rules for Stable Cathode in Lithium-Ion Battery. *Nat. Commun.* **2019**, *10*, 1650.
(123) Li, T.; Yuan, X.-Z.; Zhang, L.; Song, D.; Shi, K.; Bock, C. Degradation Mechanisms and Mitigation Strategies of Nickel-Rich NMC-Based Lithium-Ion Batteries. *Electrochem. Energy Rev.* **2020**, *3*, 43−80.
(124) Pryor, A.; Yang, Y.; Rana, A.; Gallagher-Jones, M.; Zhou, J.; Lo, Y. H.; Melinte, G.; Chiu, W.; Rodriguez, J. A.; Miao, J. GENFIRE: A Generalized Fourier Iterative Reconstruction Algorithm for High-Resolution 3D Imaging. *Sci. Rep.* **2017**, *7*, 10409.
(125) Gu, M.; Belharouak, I.; Genc, A.; Wang, Z.; Wang, D.; Amine, K.; Gao, F.; Zhou, G.; Thevuthasan, S.; Baer, D. R.; Zhang, J.-G.; Browning, N. D.; Liu, J.; Wang, C. Conflicting Roles of Nickel in Controlling Cathode Performance in Lithium Ion Batteries. *Nano Lett.* **2012**, *12*, 5186−5191.
(126) Genc, A.; Kovarik, L.; Gu, M.; Cheng, H.; Plachinda, P.; Pullan, L.; Freitag, B.; Wang, C. XEDS STEM Tomography for 3D Chemical Characterization of Nanoscale Particles. *Ultramicroscopy* **2013**, *131*, 24−32.
(127) Devaraj, A.; Gu, M.; Colby, R.; Yan, P.; Wang, C. M.; Zheng, J. M.; Xiao, J.; Genc, A.; Zhang, J. G.; Belharouak, I.; Wang, D.; Amine, K.; Thevuthasan, S. Visualizing Nanoscale 3D Compositional Fluctuation of Lithium in Advanced Lithium-Ion Battery Cathodes. *Nat. Commun.* **2015**, *6*, 8014.
(128) Wen, R.; Hong, M.; Byon, H. R. In Situ AFM Imaging of Li-O2 Electrochemical Reaction on Highly Oriented Pyrolytic Graphite with Ether-Based Electrolyte. *J. Am. Chem. Soc.* **2013**, *135*, 10870−10876.
(129) Hong, M.; Yang, C.; Wong, R. A.; Nakao, A.; Choi, H. C.; Byon, H. R. Determining the Facile Routes for Oxygen Evolution Reaction by in Situ Probing of Li-O2 Cells with Conformal Li2O2 Films. *J. Am. Chem. Soc.* **2018**, *140*, 6190−6193.
(130) Shen, Z.-Z.; Lang, S.-Y.; Shi, Y.; Ma, J.-M.; Wen, R.; Wan, L.-J. Revealing the Surface Effect of the Soluble Catalyst on Oxygen Reduction/Evolution in Li-O2 Batteries. *J. Am. Chem. Soc.* **2019**, *141*, 6900−6905.
(131) Seo, H. K.; Park, J. Y.; Chang, J. H.; Dae, K. S.; Noh, M.-S.; Kim, S.-S.; Kang, C.-Y.; Zhao, K.; Kim, S.; Yuk, J. M. Strong Stress-Composition Coupling in Lithium Alloy Nanoparticles. *Nat. Commun.* **2019**, *10*, 3428.
(132) Seo, H. K.; Hwa, Y.; Chang, J. H.; Park, J. Y.; Lee, J. S.; Park, J.; Cairns, E. J.; Yuk, J. M. Direct Visualization of Lithium Polysulfides and Their Suppression in Liquid Electrolyte. *Nano Lett.* **2020**, *20*, 2080−2086.
(133) Spurgeon, S. R.; Ophus, C.; Jones, L.; Petford-Long, A.; Kalinin, S. V.; Olszta, M. J.; Dunin-Borkowski, R. E.; Salmon, N.; Hattar, K.; Yang, W.-C. D.; Sharma, R.; Du, Y.; Chiaramonti, A.; Zheng, H.; Buck, E. C.; Kovarik, L.; Penn, R. L.; Li, D.; Zhang, X.; Murayama, M.; Taheri, M. L. et al. Towards Data-Driven Next-Generation Transmission Electron Microscopy. *Nat. Mater.* **2020**. DOI: 10.1038/s41563-020-00833-z.
(134) Bruno, I.; Gražulis, S.; Helliwell, J. R.; Kabekkodu, S. N.; McMahon, B.; Westbrook, J. Crystallography and Databases. *Data Sci. J.* **2017**, *16*, 38.
(135) Yan, H.; Voorhees, P. W.; Xin, H. L. Nanoscale X-ray and Electron Tomography. *MRS Bull.* **2020**, *45*, 264−271.
(136) Stan, T.; Thompson, Z. T.; Voorhees, P. W. Optimizing Convolutional Neural Networks to Perform Semantic Segmentation on Large Materials Imaging Datasets: X-Ray Tomography and Serial Sectioning. *Mater. Charact.* **2020**, *160*, 110119.
(137) DeCost, B. L.; Francis, T.; Holm, E. A. Exploring the Microstructure Manifold: Image Texture Representations Applied to Ultrahigh Carbon Steel Microstructures. *Acta Mater.* **2017**, *133*, 30−40.
(138) Smith, T. M.; Bonacuse, P.; Sosa, J.; Kulis, M.; Evans, L. A Quantifiable and Automated Volume Fraction Characterization Technique for Secondary and Tertiary Γ′ Precipitates in Ni-Based Superalloys. *Mater. Charact.* **2018**, *140*, 86−94.
(139) Jiang, Z.; Li, J.; Yang, Y.; Mu, L.; Wei, C.; Yu, X.; Pianetta, P.; Zhao, K.; Cloetens, P.; Lin, F.; Liu, Y. Machine-Learning-Revealed Statistics of the Particle-Carbon/Binder Detachment in Lithium-Ion Battery Cathodes. *Nat. Commun.* **2020**, *11*, 2310.
(140) Kaufmann, K.; Zhu, C.; Rosengarten, A. S.; Maryanovsky, D.; Wang, H.; Vecchio, K. S. Phase Mapping in EBSD Using Convolutional Neural Networks. *Microsc. Microanal.* **2020**, *26*, 458−468.
(141) Kunselman, C.; Attari, V.; McClenny, L.; Braga-Neto, U.; Arroyave, R. Semi-Supervised Learning Approaches to Class Assignment in Ambiguous Microstructures. *Acta Mater.* **2020**, *188*, 49−62.
(142) Liu, Z.; Bicer, T.; Kettimuthu, R.; Gursoy, D.; De Carlo, F.; Foster, I. TomoGAN: Low-Dose Synchrotron X-Ray Tomography with Generative Adversarial Networks: Discussion. *J. Opt. Soc. Am. A* **2020**, *37*, 422−434.
(143) Yang, X.; De Carlo, F.; Phatak, C.; Gursoy, D. A Convolutional Neural Network Approach to Calibrating the Rotation Axis for X-Ray Computed Tomography. *J. Synchrotron Radiat.* **2017**, *24*, 469−475.
(144) Young, S. R.; Maksov, A.; Ziatdinov, M.; Cao, Y.; Burch, M.; Balachandran, J.; Li, L.; Somnath, S.; Patton, R. M.; Kalinin, S. V.; Vasudevan, R. K. Data Mining for Better Material Synthesis: The Case of Pulsed Laser Deposition of Complex Oxides. *J. Appl. Phys.* **2018**, *123*, 115303.
(145) Shi, Y.; Zhang, M.; Fang, C.; Meng, Y. S. Urea-Based Hydrothermal Synthesis of Lini0.5Co0.2Mn0.3O2 Cathode Material for Li-Ion Battery. *J. Power Sources* **2018**, *394*, 114−121.
(146) Lee, S.-H.; Lee, S.; Jin, B.-S.; Kim, H.-S. Optimized Electrochemical Performance of Ni Rich LiNi0.91Co0.06Mn0.03O2 Cathodes for High-Energy Lithium Ion Batteries. *Sci. Rep.* **2019**, *9*, 8901.
(147) Jang, J. H.; Kim, Y.-M.; He, Q.; Mishra, R.; Qiao, L.; Biegalski, M. D.; Lupini, A. R.; Pantelides, S. T.; Pennycook, S. J.; Kalinin, S. V.; Borisevich, A. Y. In Situ Observation of Oxygen Vacancy Dynamics and Ordering in the Epitaxial LaCoO3 System. *ACS Nano* **2017**, *11*, 6942−6949.
(148) Kotakoski, J.; Mangler, C.; Meyer, J. C. Imaging Atomic-Level Random Walk of a Point Defect in Graphene. *Nat. Commun.* **2014**, *5*, 3991.
(149) Ishikawa, R.; Mishra, R.; Lupini, A. R.; Findlay, S. D.; Taniguchi, T.; Pantelides, S. T.; Pennycook, S. J. Direct Observation of Dopant Atom Diffusion in a Bulk Semiconductor Crystal Enhanced by a Large Size Mismatch. *Phys. Rev. Lett.* **2014**, *113*, 155501.
(150) Dyck, O.; Kim, S.; Jimenez-Izal, E.; Alexandrova, A. N.; Kalinin, S. V.; Jesse, S. Building Structures Atom by Atom via Electron Beam Manipulation. *Small* **2018**, *14*, 1801771.
(151) Horwath, J. P.; Zakharov, D. N.; Mégret, R.; Stach, E. A. Understanding Important Features of Deep Learning Models for Segmentation of High-Resolution Transmission Electron Microscopy Images. *npj Comput. Mater.* **2020**, *6*, 108.
(152) Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X.; Sang, X.; Xiao, K.; Unocic, R. R.; Vasudevan, R.; Jesse, S.; Kalinin, S. V. Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. *ACS Nano* **2017**, *11*, 12742−12752.
(153) Li, W.; Field, K. G.; Morgan, D. Automated Defect Analysis in Electron Microscopic Images. *npj Comput. Mater.* **2018**, *4*, 36.

(154) Ziletti, A.; Kumar, D.; Scheffler, M.; Ghiringhelli, L. M. Insightful Classification of Crystal Structures using Deep Learning. *Nat. Commun.* **2018**, *9*, 2775.
(155) Aguiar, J. A.; Gong, M. L.; Unocic, R. R.; Tasdizen, T.; Miller, B. D. Decoding Crystallography from High-Resolution Electron Imaging and Diffraction Datasets with Deep Learning. *Sci. Adv.* **2019**, *5*, No. eaaw1949.
(156) Kaufmann, K.; Zhu, C.; Rosengarten, A. S.; Maryanovsky, D.; Harrington, T. J.; Marin, E.; Vecchio, K. S. Crystal Symmetry Determination in Electron Diffraction using Machine Learning. *Science* **2020**, *367*, 564.
(157) Yao, L.; Ou, Z.; Luo, B.; Xu, C.; Chen, Q. Machine Learning to Reveal Nanoparticle Dynamics from Liquid-Phase TEM Videos. *ACS Cent. Sci.* **2020**, *6*, 1421−1430.
(158) Belianinov, A.; Vasudevan, R.; Strelcov, E.; Steed, C.; Yang, S. M.; Tselev, A.; Jesse, S.; Biegalski, M.; Shipman, G.; Symons, C.; Borisevich, A.; Archibald, R.; Kalinin, S. Big Data and Deep Data in Scanning and Electron Microscopies: Deriving Functionality from Multidimensional Data Sets. *Adv. Struct. Chem. Imaging* **2015**, *1*, 6.
(159) Agar, J. C.; Naul, B.; Pandya, S.; van der Walt, S.; Maher, J.; Ren, Y.; Chen, L.-Q.; Kalinin, S. V.; Vasudevan, R. K.; Cao, Y.; Bloom, J. S.; Martin, L. W. Revealing Ferroelectric Switching Character using Deep Recurrent Neural Networks. *Nat. Commun.* **2019**, *10*, 4809.
(160) Holstad, T. S.; Ræder, T. M.; Evans, D. M.; Småbråten, D. R.; Krohns, S.; Schaab, J.; Yan, Z.; Bourret, E.; van Helvoort, A. T. J.; Grande, T.; Selbach, S. M.; Agar, J. C.; Meier, D. Application of a Long Short-Term Memory for Deconvoluting Conductance Contributions at Charged Ferroelectric Domain Walls. *npj Comput. Mater.* **2020**, *6*, 163.
(161) Frey, N. C.; Akinwande, D.; Jariwala, D.; Shenoy, V. B. Machine Learning-Enabled Design of Point Defects in 2D Materials for Quantum and Neuromorphic Information Processing. *ACS Nano* **2020**, *14*, 13406−13417.
(162) Wu, S.; Kondo, Y.; Kakimoto, M.-A.; Yang, B.; Yamada, H.; Kuwajima, I.; Lambard, G.; Hongo, K.; Xu, Y.; Shiomi, J.; Schick, C.; Morikawa, J.; Yoshida, R. Machine-Learning-Assisted Discovery of Polymers with High Thermal Conductivity using a Molecular Design Algorithm. *npj Comput. Mater.* **2019**, *5*, 66.
(163) Court, C. J.; Yildirim, B.; Jain, A.; Cole, J. M. 3-D Inorganic Crystal Structure Generation and Property Prediction via Representation Learning. *J. Chem. Inf. Model.* **2020**, *60*, 4518−4535.
(164) Ma, B.; Wei, X.; Liu, C.; Ban, X.; Huang, H.; Wang, H.; Xue, W.; Wu, S.; Gao, M.; Shen, Q.; Mukeshimana, M.; Abuassba, A. O.; Shen, H.; Su, Y. Data Augmentation in Microscopic Images for Material Data Mining. *npj Comput. Mater.* **2020**, *6*, 125.
(165) Ohno, H. Training Data Augmentation: An Empirical Study using Generative Adversarial Net-Based Approach with Normalizing Flow Models for Materials Informatics. *Appl. Soft. Comput.* **2020**, *86*, 105932.
(166) Ohno, H. Auto-Encoder-Based Generative Models for Data Augmentation on Regression Problems. *Soft Comput.* **2020**, *24*, 7999−8009.
(167) Agar, J. C.; Cao, Y.; Naul, B.; Pandya, S.; van der Walt, S.; Luo, A. I.; Maher, J. T.; Balke, N.; Jesse, S.; Kalinin, S. V.; Vasudevan, R. K.; Martin, L. W. Machine Detection of Enhanced Electromechanical Energy Conversion in PbZr0.2Ti0.8O3 Thin Films. *Adv. Mater.* **2018**, *30*, 1800701.
(168) Agar, J. C.; Naul, B.; Pandya, S.; van der Walt, S.; Maher, J.; Ren, Y.; Chen, L.-Q.; Kalinin, S. V.; Vasudevan, R. K.; Cao, Y.; Bloom, J. S.; Martin, L. W. Revealing Ferroelectric Switching Character using Deep Recurrent Neural Networks. *Nat. Commun.* **2019**, *10*, 4809.
(169) Vasudevan, R. K.; Ziatdinov, M.; Vlcek, L.; Morozovska, A. N.; Eliseev, E. A.; Yang, S.-Z.; Gong, Y.; Ajayan, P.; Zhou, W.; Chisholm, M. F.; Kalinin, S. V. Investigating Phase Transitions from Local Crystallographic Analysis Based on Machine Learning of Atomic Environments. *arXiv (Condens. Matter)*, 2020. arXiv:2005.10001.https://arxiv.org/abs/2006.10001.
(170) Groschner, C.; Choi, C.; Nguyen, D.; Ophus, C.; Scott, M. Machine Learning for High Throughput HRTEM Analysis. *Microsc. Microanal.* **2019**, *25*, 150−151.
(171) Vasudevan, R. K.; Kelley, K.; Funakubo, H.; Jesse, S.; Kalinin, S. V.; Ziatdinov, M. Autonomous Experiments in Scanning Probe Microscopy and Spectroscopy: Choosing Where to Explore Polarization Dynamics in Ferroelectrics. *arXiv (Condens. Matter)*, 2020. arXiv:2011.13050. https://arxiv.org/abs/2011.13050.
(172) Burgess, C. P.; Higgins, I.; Pal, A.; Matthey, L.; Watters, N.; Desjardins, G.; Lerchner, A. Understanding Disentangling in β-VAE. *arXiv (Machine Learning: Statistics)*, 2018. arXiv:1804.03599. https://arxiv.org/abs/1804.03599.
(173) Böhm, V.; Seljak, U. Probabilistic Auto-Encoder. *arXiv (Machine Learning: Computer Science)*, 2020. arXiv:2006.05479. https://arxiv.org/abs/2006.05479.
(174) Champion, K.; Lusch, B.; Kutz, J. N.; Brunton, S. L. Data-Driven Discovery of Coordinates and Governing Equations. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 22445.
(175) Duarte, J.; Han, S.; Harris, P.; Jindariani, S.; Kreinar, E.; Kreis, B.; Ngadiuba, J.; Pierini, M.; Rivera, R.; Tran, N.; Wu, Z. Fast Inference of Deep Neural Networks in Fpgas for Particle Physics. *J. Instrum.* **2018**, *13*, P07027.
(176) Draxl, C.; Scheffler, M. NOMAD: The FAIR Concept for Big Data-Driven Materials Science. *MRS Bull.* **2018**, *43*, 676−682.
(177) Aykol, M.; Hummelshøj, J. S.; Anapolsky, A.; Aoyagi, K.; Bazant, M. Z.; Bligaard, T.; Braatz, R. D.; Broderick, S.; Cogswell, D.; Dagdelen, J.; Drisdell, W.; Garcia, E.; Garikipati, K.; Gavini, V.; Gent, W. E.; Giordano, L.; Gomes, C. P.; Gomez-Bombarelli, R.; Balaji Gopal, C.; Gregoire, J. M.; Grossman, J. C.; Herring, P.; Hung, L.; Jaramillo, T. F.; King, L.; Kwon, H.-K.; Maekawa, R.; Minor, A. M.; Montoya, J. H.; Mueller, T.; Ophus, C.; Rajan, K.; Ramprasad, R.; Rohr, B.; Schweigert, D.; Shao-Horn, Y.; Suga, Y.; Suram, S. K.; Viswanathan, V.; Whitacre, J. F.; Willard, A. P.; Wodo, O.; Wolverton, C.; Storey, B. D.; et al. The Materials Research Platform: Defining the Requirements from User Stories. *Matter* **2019**, *1*, 1433−1438.

https://dx.doi.org/10.1021/acsnano.1c00211
ACS Nano 2021, 15, 3971−3995