Report of the Workshop: Scientific Opportunities and Instrumentation Needs for Next Generation Materials Genomics Based Materials Research in Materials with Long-Range Order

1 Organization

1.1 Context

This is a report from a workshop and information gathering exercise on Materials Science Laboratories of the Future (LoF). The workshop was held virtually on 16-17th November 2022 and was one of a series of three “Materials Laboratories of the Future: Instrumentation and Infrastructure to Accelerate the Unification of the Materials Innovation Infrastructure” workshops sponsored by the National Science Foundation (NSF). This workshop focused on materials with long-range order, such as bulk crystals, epitaxial films, two-dimensional materials, Van der Waals solids, and other materials science topics. Prior to the workshop, the Materials Science community was polled through a questionnaire that was distributed widely to organizations and departments in the U.S.

1.2. Modality of the meeting and questionnaire

The goal of the meeting was to obtain broad input from the materials science community about scientific needs, critical gaps in experimental capabilities, and visions for LoFs that could help to address the gaps, specifically focusing on “Materials with Long Range Order (MLRO)”. Input was also sought to identify training and human resource requirements that will arise in LoFs as well as opportunities for broader participation that will be enabled by LoFs.

1.3 People:

1.3.1 Chairs

Simon J. L. Billinge, Professor of Materials Science and of Applied Physics and of Applied Mathematics, Columbia University
John Mitchell, Senior Scientist, Materials Science Division, Argonne National Laboratory
Susanne Stemmer, Professor of Materials, University of California, Santa Barbara

1.3.2 Advisory Committee

Dmitri Basov, Professor of Physics, Columbia University
David Goldhaber-Gordon, Professor of Physics, Stanford University
Tyrel McQueen, Professor of Chemistry, Johns Hopkins University
Daniel Lopez, Professor of Electrical Engineering and Computer Science, Penn State University

2 Introduction

The field of Materials Science and Engineering (MSE) encompasses fundamental science and technological developments in materials that are the foundation for solutions to many of the most challenging of human problems such as sustainability, energy, health, information sciences, defense, and environmental protection. In the US, experimental MSE research is carried out in laboratories at universities, national laboratories, and in industry. The bulk of fundamental experimental MSE research, which lays the foundations of future transformative technologies, is carried out in the laboratories of individual investigators in academia and national laboratories. In some areas of MSE research, centralized large-scale user facilities, such as high-power x-ray and neutron sources, and high-field magnets, play an important role. Many areas for MSE rely on an infrastructure of experimental facilities that are exceptionally expensive to equip and maintain, such as nanofabrication facilities (“clean rooms”), synthesis equipment, or advanced microscopes. These “laboratories of today” (LoT) are the central incubators of scientific discoveries, new knowledge in materials, and their implementation into advanced technologies. These LoT have also trained generations of scientists and engineers in advanced skills that are essential to the nation’s industry, improving people’s lives and powering the US economy.

The materials LoT face an increasing number of challenges. In particular, the nation’s universities have enormous difficulties in equipping LoT with state-of-the-art MSE instrumentation and in maintaining the advanced equipment. This is primarily due to the lack of extramural support for instrumentation, infrastructure, and technical staff. Even the nation’s top research universities are falling behind their peers in competitor countries in their ability to provide the necessary up to date infrastructure and advanced instrumentation for MSE, with potentially dire consequences for future innovation in MSE and US leadership in this endeavor.

However, the focus of this workshop was not on the challenges facing the nation's LoT. Rather, the workshop was focused on identifying the potential for transformational changes in our experimental capabilities and approaches, as implemented in LoF. A particular emphasis was on gathering input from the community with regards to potential transformational changes that may be enabled by harnessing emerging capabilities in automation, data analytics, machine learning (ML), and artificial intelligence (AI), to empower scientists and engineers with new tools and approaches to research. The Materials Genome Initiative strategic plan [1] offers the prospect that advanced AI/ML and data analytical approaches can accelerate discovery of new materials. Exploring the potential of incorporating such approaches in an experimental MSE LoF is timely, given the rapidly emerging capabilities in these fields. However, it is important to emphasize that these data-driven approaches are by no means the exclusive pathway to the LoF and the community was also asked to provide a broader perspective on their vision of LoF.

Here, with input from the broad materials science community, we identify examples of challenging questions at the forefront of MSE as applied to materials with long range order, discuss the critical gaps and ways to bridge them, and express a vision for a path to LoF. We identify emerging possibilities, but also challenges, associated with developments such as AI/ML. We also consider transformative LoF developments that are not genomic in nature.

2.1 Findings

2.1.1 Critical research needs and capability gaps

The workshop explored the vision and needs for LoF in these four MRLO sub-areas:

1. Quantum device creation, quantum state characterization
2. Engineering and characterizing functional devices
3. Predictive synthesis with control and characterization at the level of individual atoms and defects
4. Probing and understanding the dynamics of atoms and electrons in MRLO

Results of the discussions are presented in detail in Sections 3 (science) and 4 (research, training, and access) of this report. In this summary section we highlight findings on the needs and gaps that cut across more than one of the sub-areas mentioned above. Addressing these gaps can be expected to have significant impacts.

There is a critical need for revolutionary new experimental tools and capabilities:

- That increase by orders of magnitude the throughput in synthesis of film, bulk and low-dimensional materials

- That transform our capabilities for *in situ* monitoring of structure and properties during growth/synthesis
- That substantially increase (by up to orders of magnitude) the sensitivity of quantitative methods for detecting and characterizing defects, surfaces and interfaces in materials
- That improve spatially and temporally resolved measurements of materials properties and that can probe at the length/time scales of relevant defects and device structures
- That rapidly fabricate electronic and quantum devices from materials having any shape or dimension – bulk, individual grains, thin films, whiskers, etc.
- That effectively capture, with low-overhead and low-cost, accurate data and metadata from experiments for professionally maintained FAIR databases, and that will reduce the burden on individual PIs and institutions to accomplish this task

At the same time, significant gaps in scientific data and simulation must also be bridged:

- There is still a critical gap between the outputs of simulations and the inputs of experiments. The quantities that theory computes are frequently not the same quantities that experimentalists measure or control as inputs. This inhibits effective closed-loop theory - simulation - experiment protocols.
- There is a critical gap between the community organized, professionally maintained and curated, and sustainably financed FAIR data infrastructure that is needed and today's increasingly confusing landscape of *ad hoc* databases of variable quality, objectives, and sustainability. The obligation to manage, curate, and distribute data, whilst being responsible for privacy protocols and protecting intellectual property places an unsustainable burden on individual PIs.
- There is a lack of support for assisting and training of the MSE community in matters of data provenance, and to correctly incentivize the collection, sharing and value extraction from experimental data

A successful vision and strategy of a LoF must provide the means to create these new tools and capabilities and to close these critical gaps.

2.1.2 Visions and challenges for laboratories of the future

At the heart of the workshop discussions was an exploration of a diverse set of visions for LoF. **A significant finding of the workshop was that while the needs of LoT are generally well-understood, many of the visions for LoF, and how they would serve to address the critical needs and capability gaps identified, are speculative at best. Formulating a clear vision for LoF will continue to require**

interrogating, prototyping, and assessing the effectiveness of the various MGI driven approaches to laboratory experimentation and instrumentation.

At its heart, MSE is an experimental enterprise with an immense need for highly sophisticated, adaptable, laboratory-scale, instrumentation suitable for exploratory research. In this regard, MSE research differs fundamentally from manufacturing, where great productivity gains have been made by automation, allowing robots to execute repetitive tasks efficiently and with very high reproducibility. Historically, leaps in scientific discovery and productivity in research follow a very different path. Discoveries emerge from nonlinear, out-of-the box creative thinking and the human ability to invent, rather than from optimizing existing approaches.

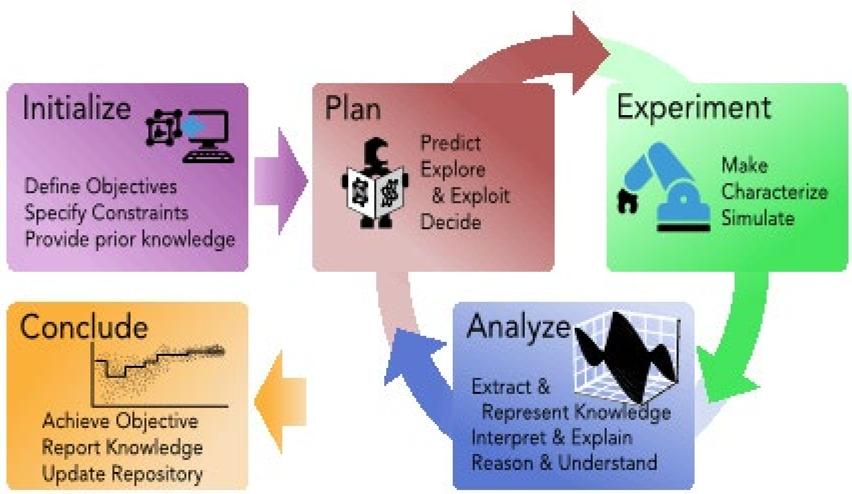
Consider the following breakthroughs made in MRLO in the past decades. High-temperature superconductors were discovered *because* scientists went against the prevailing paradigm of the day by exploring materials that almost no-one believed to hold any promise whatsoever for exhibiting superconductivity. The discovery of topology as a new guiding principle in condensed matter was an astounding intellectual feat that led to an explosion of materials discoveries with new properties. Some of the most exciting discoveries in MRLO were completely unexpected and often revealed in a single experiment, rather than through a search for something expected or predicted. An example for this is the discovery of the fractional quantum Hall effect. It is these “rare events” that have proved transformative for MSE over and over again. All these discoveries were made by leaps outside of the (at the time) known materials phase space and thus not readily accessible to ML/AI-type of approaches that learn from prior knowledge. These examples make it clear that LoF must preserve the capacity to support and enhance this distinctly human endeavor of scientific discovery.

At the same time, there is a demonstrable need for improving throughput and output, for reducing failure for “trivial” reasons that do not contribute to understanding, and for greater efficiency, reproducibility, and robustness in experimental MSE. These areas are obvious targets for LoF. The beginnings of a transformation are already here. Sophisticated, adaptive approaches to laboratory automation are emerging, which we refer to as “autonomation.” Autonomation allows computer controlled, automated processes (robots) to carry out distinctly non-repetitive things, such as adapting their workflows in smart ways to respond to changing experimental outputs. At present, demonstrations of such “autonomated” experiments are in a formative stage, for example, smart optimizations over a high dimensional input parameter space of repetitive tasks. Even this early stage, this type of LoF may benefit many materials science problems, and current activities are harvesting the low hanging fruit in this space.

But clearly the goal of an autonomous LoF would be to go well beyond this emerging autonomation, to aspire to *undirected* experimentation, powered by data and guided by artificial intelligence. We may call this “autonomously aided discovery” (AAD) where the AI and the human work together each doing the tasks they do best. What might

LoF look like within this AAD framework? The answer is unclear. We do not know what is possible; we do not know what is necessary to accomplish it; we do not know how to harness its potential; we do not even know what this full potential is; and we do not know its limitations and pitfalls. AAD is thus, itself, a research grade problem.

An example of research in AAD is the Air Force Research Laboratory's ARES™, a closed-loop autonomous research system for materials development [2]. A rapidly growing number of researchers are now building their own autonomous research systems, exploiting advances in artificial intelligence (AI), autonomy and robotics, along with modeling and simulation, to create research robots capable of doing iterative experimentation orders of magnitude faster than today.



A schematic of the autonomous experiment workflow. Scientific goals and objectives are specified by a human who initiates the autonomous loop which is designed to achieve those goals. A convergence criterion causes the loop to be exited. Results are presented back to the human researcher during iteration and after convergence for interpretation and conclusion drawing [2].

To progress, what is needed today is an effort to explore the realm of possibilities with prototype setups, the results of which can articulate a path towards materials LoF. This path could begin by exploring "experiments of the future" (EoF), which would offer exciting demonstrations of capabilities. Standing alone, EoF will deliver meaningful scientific and technological impact. Collectively, these EoF will become the building blocks of the LoF. *Research and infrastructure investments into EoF should be a high priority to maintain leadership and allow US researchers to realize the gains that we* *believe will be possible with such developments*. Intermediate stops on the journey could look like incrementally complex “smart assistants” that provide context for human researchers to make creative decisions more quickly, efficiently, and based on greater and more precise intelligence.

Proper automation can enable remote operation which we refer to as “Measurements as a service” (MaaS), using as an analogy software as a service (SaaS), resulting in broader access to routine measurements.

Autonomous experimentation is built on experimental hardware that can be automated into complex autonomous workflows that link one instrument to another. Advances are needed in the decision methods themselves for AAD where the processes are characterized by an inherent high iteration rate, but potentially low data rates and the need to incorporate physics-informed constraints. To deliver on its promise, future objectives for autonomous experimentation should aim at explicit hypothesis search and testing, resulting in useful exploitable scientific knowledge, rather than blackbox statistical representations that can result from naïve (i.e., no physical representation) machine learning approaches.

The pathway to this stage of autonomously aided discovery in the LoF will require significant funding not only for research but also for data infrastructures. Data capture and dissemination efforts in experimental materials science are currently *ad hoc*. Many open questions remain and need to be addressed, such as transferability of experimental data and metadata between laboratories (e.g., synthesis protocols), the quality of the data, individual privacy, and intellectual property concerns. Developing and maintaining a robust, contemporary, curated data infrastructure in MSE lies far beyond the scope of most individual PIs and their students.

The development of EoF and LoF requires simultaneous attention to potential pitfalls. Innovation will likely not benefit from ever increasing numbers of incremental advances made by automation, generation of vast amounts of data of questionable quality or utility, or by simply increasing throughput, productivity, and output. These should not be the metrics by which advances in science, and therefore also EoF and LoF, are judged. The pitfalls of AI/MG based approaches of getting trapped in the “known” universe or, worse, within prevailing but wrong paradigms, are well documented.

There are significant concerns regarding how the quality of data in databases can be assessed and to whom this burden will fall. At present, quality control in science rests on individuals investigating the reproducibility of results of experiments reported by others. This needs to be enhanced with rich data capture protocols. However, in a world where widespread capture and storage of accurate data and metadata is the standard, great care needs to be taken to develop no blame cultures and to address concerns of privacy. A fear of mistakes that are easily attributable because of the rich contextual data and metadata capture can lead to a culture of risk avoidance that suppresses discovery, as well as causing harm to individuals’ reputations or careers.

The LoF is a decadal research project in its own right, a grand challenge that will demand persistence, innovation, and sustained support.

2.2 Community-Recommended areas of focus

Key outcomes of the workshop were suggestions of focus areas that could form the basis of a strategy for pursuing research into AAD and delivering the promise of LoF:

- Reestablishing US leadership in materials science research infrastructure through renewed investment in modern MSE laboratory equipment and the required infrastructure at universities.
- Sustaining state-of-the-art laboratory equipment in universities and national laboratories with long-term technical staff support.
- Creating a balanced portfolio between national and regional shared facilities and in individual PI labs.
- Supporting high-risk, high-reward interdisciplinary research projects into prototype EoF.
- Developing a sustainable model for shared FAIR databases based on community standards for data dictionaries and schemas.
- Managing privacy concerns and developing no blame cultures in a data-rich workplace.
- Exploring instrument access models that broaden participation.
- Balancing the needs of easy, low activation-barrier access to instrumentation to train the next generation of methods developers with deep understanding of the underlying instrumentation and science.

3. Scientific and technological challenges and opportunities in materials with LRO

3.1 Quantum Materials and Devices

Over the past decade, the potential of utilizing quantum states and phenomena for quantum information sciences, sensing, and computing has generated enormous excitement in the condensed matter and materials physics academic communities as well as in industry, not only as a fertile ground for scientific discoveries, but also because of its potential for new technologies. The discovery, development, and deployment in devices of quantum materials is central to this endeavor. For example, the experimental realization of topological matter is an essential milestone in the search for new quasi-particle states that may enable quantum computers that are more robust against “decoherence,” the decay of entangled quantum states into an incoherent, classical assembly. Decoherence remains a defining challenge in the experimental realization and scale-up of quantum devices and systems.

Synthesis of new quantum materials and exploration of their unique states now takes place in many academic laboratories across the world. Many of these efforts are a direct consequence of strategic investments that recognized critical needs. One success story emerged from an influential report, "*Frontiers in Crystalline Matter. From Discovery to Technology*", published in 2009 by the National Academies [3], which led to a reinvigoration in the US in areas of discovery and growth of novel crystalline materials as single crystals. Once a rarity, there are now many academic laboratories in the U.S. with the ability to grow large single crystals of a wide range of materials. Similarly, the Designing Materials to Revolutionize and Engineer Our Future (DMREF) program [4] of the National Science Foundation has focused on supporting the discovery of new materials through a feedback loop of computational searches, synthesis and characterization. Without doubt, these initiatives have been enormously successful. To name a few examples, stable topological materials, of which once only a handful were known, are now considered ubiquitous [5-7]. Similarly, following the discovery of graphene, atomically-thin materials have now become an entire academic sub-field of their own, and shown to have many rich and unique properties not found in other materials classes.

Given a wealth of new quantum materials that have become available to researchers and industry over the past decades, it is fair to ask why more of these materials have not found their way into new device technologies and, specifically, what would be the role LoT and LoF play in promoting the needed translation. This workshop identified a number of capability gaps, in particular a large gap in scientific knowledge and experimental capabilities needed to deploy quantum materials into devices.

While theoretical discovery and bulk synthesis efforts of new quantum materials candidates have been thriving, experimental investment is sorely lacking into experimental capabilities that allow for assessing, characterizing, and realizing even the most essential building blocks needed for enabling solid-state quantum information systems with realistic device structures within these materials. Equally importantly, the science of such quantum devices, for example, how to perform the necessary operations for a particular application, has often not yet been worked out. Given that there are wide remaining gaps in scientific knowledge, we are rather far from the point where tasks like scaling or optimization are relevant. There is significant industrial research in particular materials systems for quantum information which may benefit from current autonomous processes. However, this workshop views it as essential to invest in academic research, at a level on a par with industrial investments, because the latter are unlikely to address the large scientific gaps. For example, to date, experimental techniques that can assess and witness entanglement and fractionalized excitations in solid-state systems are largely unknown, as are techniques that can fingerprint the role of specific material imperfections in the decoherence process in qubits.

Experimental realization and testing of solid-state quantum information systems requires highly advanced and very expensive infrastructure that is comparable with what is used in the most advanced semiconductor technologies. For example, fabricating a qubit from a topological material, which has not yet been demonstrated, requires thin film epitaxy, nanofabrication, and ultra-low temperature characterization facilities. Unfortunately, existing LoT that specialize in semiconductor and nanodevice fabrication and which exist at many major research universities in the nation, can only be partially adapted to this new need. For instance, epitaxial growth of topological matter cannot be carried out in an existing molecular beam epitaxy system that is also used for conventional semiconductor thin film materials. Fabrication of superconducting qubits requires dedicated metallization chambers in nanofabrication facilities that are free of magnetic impurities that may be used in other processes taking place in the same facility. These examples make it clear that substantial investment in nanofabrication facilities specifically for quantum devices is needed for further progress.

Beyond quantum device fabrication, significant investment is needed to characterize novel quantum materials, states, and phenomena such as decoherence and entanglement. Often, the scientific foundation of such methods is not yet worked out and still a matter of debate. Support is therefore required in the development of quantum device characterization approaches. Theory could play an important role in these developments.

The realization of solid-state quantum information systems often involves integrating materials with dissimilar properties, such as superconducting materials interfaced with strong spin-orbit coupling materials to create Majorana fermion states. However, this integration is only partially achieved even at the leading-edge academic research level, and achieving it requires further improvement or novel designs for epitaxy equipment that can create currently unachievable heterostructures of dissimilar materials. In addition, the quantum phenomena found in such heterostructures are often confined to a few atomic layers near an interface and characterizing and optimizing such interfacial phenomena requires advances in techniques to identify the structural, electronic, and chemical properties of the buried interface. For example, incident energy-dependent resonant x-ray scattering or reflectometry could be further developed to improve the characterization of the interfacial phenomena.

AI/ML-based approaches may play a role in aspects of optimization and improving throughput of experiments, particularly those that involve repetitive tasks. In general, given their complexity and the involvement of many fabrication steps, the failure rate of a quantum device experiment from start (materials synthesis) to finish (device measurement) is typically very high, and, as a result, data is sparse. If throughput could be increased, failure rates decreased, and successful and failed experimental metadata and data effectively captured, scientific progress would be accelerated.

We emphasize the need for progress in the scientific underpinnings of quantum devices. Many of the potentially transformative capabilities of quantum technologies are far from deployment-ready. Thus, both long-term investment in scientists and the experimental facilities to test and realize their ideas are needed to transform the nation's investments in quantum materials into technological successes.

3.2 Atomic-scale devices

The anticipated potential of nanoscience and nanotechnology will only be realized if atomic-scale structures can be precisely and reproducibly fabricated and assembled into a working nanosystem that can be probed, controlled, and manipulated in a reliable manner [8]. The most significant obstacles to the implementation of functional nanosystems come from limitations associated with the reproducibility of existing nanofabrication processes [9], the design and integration of atomic and nanoscale materials and structures [10], the lack of understanding of the fundamental forces dominant at the nanoscale [11], and the limited existing capabilities for manipulation and probing of nanostructures [12].

The promises and challenges of nanomaterial research are at the center of a wide array of central questions in materials science, and it has been the focus of the National Nanotechnology initiative for the past 23 years because of its broad, cross-cutting nature. Urgent current scientific examples where nanoscience is expected to play a pivotal role have been highlighted in recent workshops and reports such as DOE-BES's 2017 *Basic research needs workshop for catalysis science* and *Basic research needs workshop on next generation electrical energy storage* [13], among multiple others, and NSF goals in *Nanoscale structures, novel phenomena and quantum control*, *Device and system architecture* (integrating nanoscale devices into measurement and control assemblies), *Multi-scale, multi-phenomena modeling and simulation at the nanoscale.* A relevant theme in the context of the current report is *Converging technologies* (the convergence of nanotechnology with information technology, biology and the social sciences).

Nanoscale and atomic scale physics also has a role to play in the emerging field of quantum information science (QIS) discussed above. Challenges exist in the fabrication and probing of qubits, and quantum device engineers are trying to build them [14] using industrial semiconductor manufacturing processes. Investments are needed to accelerate this task. Also, the nascent area of quantum sensing requires a subtle balance between coupling the sensor to the environment while maintaining the quantum state, a critical condition that current packaging architectures cannot deliver [15].

The properties of slender films, particularly atomically thin films, is an active area of research [16-19] that impacts many scientific areas such as biology [20], flexible electronics [21], condensed matter physics [22], organ engineering [23] and mechanical metamaterials [24]. Multiple scientific groups are working toward understanding the fundamental science behind the mechanical properties of slender films, but the lack of reproducibility in the fabrication processes, characterization standards, and comprehensive theoretical methods to simulate fluctuation-induced forces at the nanoscale, are critical barriers to reaching a complete understanding [25,26] of these unique phenomena.

Mechanical oscillators are an essential component of practically every electronic device requiring a frequency reference for timekeeping or synchronization [27] and are also widely used in frequency-shift-based sensors of mass [28], force [29] and magnetic field [30]. As the dimensions of the vibrating structures are reduced from the micro- to the atomic scale, thermal and quantum fluctuations and quenched defects become relevant, and their mechanical stability and dynamics cannot be described by either the well-known Hooke's law or fundamental scaling laws. These phenomena are especially relevant for atomically thin 2D materials, where elastic deformation and electronic degrees of freedom are closely coupled. Control and monitoring of atomic-scale devices is needed but presents the additional challenge of dealing with structures that exhibit strong nonlinearities in their mechanical response [31]. In most situations, the nonlinear response is actually caused by the inability to apply small enough forces to these systems whilst increasing the detected signal above the background noise. Sometimes, because of their small dimensions, thermal fluctuations are strong enough to drive them into the nonlinear regime [32]. Consequently, nonlinear responses in atomic-scale devices is the rule rather than the exception, and it needs to be fully understood for engineering and characterizing atomic-scale devices.

Defects (impurities, point/extended defects) play an important role in semiconductor materials and devices. Understanding and connecting the effect on properties of defects across length scales from atomic to device scale is important but unsolved. Secondary ion mass spectrometry (SIMS) has high sensitivity (down to 10^{15} cm^{-3} or lower) but lacks high spatial resolution down to nm scales. High resolution transmission electron microscopy (TEM) allows us to probe crystal imperfections with atomic resolution, but low concentrations (e.g., ppm) of impurities/defects are difficult to detect. Neutron (elastic) scattering can be performed on 10s of nm thick films, but it requires sample sizes on mm-scales.

Taken together, high throughput defect characterization with high sensitivity from atomic to device scale is a critical need and understanding the correlation between material defects and properties with device performance will require an extensive research effort. Understanding these fundamental challenges and addressing the related technical issues would provide realistic design rules for atomic-scale devices and their incorporation into new technologies.

Many approaches are currently under development to fabricate nano-scale structures and devices with atomic precision. Nanomanufacturing techniques such as atomic-scale 3D printers [33] and tip-based nanofabrication [34] are promising alternatives for growing atomic-scale devices on demand. However, they lack the reliability and reproducibility standards required to have a significant presence in R&D labs. Alternatively, the trillion-dollar semiconductor industry has been manufacturing transistors reproducibly over the past 50 years using advanced semiconductor manufacturing methods, and are now pushing the dimensions of the transistors to a few nm in size [35]. These processes achieve near atomic-scale reproducibility and uniformity by using the world's most advanced nanofabrication tools under strict operation rules.

With this in mind, establishing strong bridges between academia and industry could deliver controllable and reproducible manufacturing techniques for engineering atomic-scale devices. For example, a recent collaboration between researchers from TU Delft and Intel, demonstrated scalable and reliable manufacturing of Si metal-oxide-semiconductor qubits [36] by giving university researchers access to state-of-the-art fabrication lines at Intel.   Future public-private partnerships could be a feature of LoF.

A similar effort in the field of microelectronics is under development by NIST, Google, and Skywater [37]. These organizations have signed a cooperative research agreement to provide microelectronics chips to researchers for developing new nanotechnology devices. Similar agreements with enterprises manufacturing nanomaterials and developing metrology equipment would ensure reliable access to the technology needed to engineer atomic-scale devices.

To accelerate the development pace of future semiconductors, it is necessary to establish an integrated material characterization platform that enables high throughput defects characterization from atomic to macro scale. Bridging this gap would allow efficient mapping and correlation between defects properties and device performance. Automatic data collection, including the incorporation of AI/ML and autonomous approaches, would provide efficient feedbacks to accelerate the discovery and development of new semiconductors for future applications.

The challenges mentioned above are ubiquitous for any field of science trying to understand the fundamentals of atomic-scale materials and devices. Without accessible and reliable manufacturing processes and characterization techniques the engineering of atomic-scale devices will take decades to benefit society.

While the term "lab-to-fab" is commonly used when describing the translation of basic research into a commercial product, we believe that an effective LoF should routinely embrace the "fab-to-lab" approach. Under this paradigm, academic researchers would have access to advanced nanomanufacturing technology at a reduced cost, and universities won't need to support expensive campus cleanrooms. We believe that a "fab-to-lab" approach to atomic-scale devices manufacturing would significantly help push forward this nascent field's engineering and transformative applications.

3.3 Predictive synthesis

New materials are the lifeblood of technological innovation. Whether it is a new battery cathode material, a high performance thermoelectric, or a new photovoltaic material, the need to study, understand, and ultimately control the function of MLRO demands that we be able to make them, first in the laboratory and eventually at scale. With the growing complexity that links composition, structure, and multiscale architecture of today's and tomorrow's functional materials, both discovery synthesis and targeted synthesis of MLRO are growing ever more challenging. As such, the growing need to improve outcomes of complex materials synthesis experiments makes a compelling argument for a synthesis LoF, where the design-create-measure loop can be traversed more rapidly, more intelligently, and more predictably — or reimagined entirely.

Indeed, the predictive synthesis of materials at-will remains an unrealized 'holy grail' of the materials by design paradigm. Today, anyone with a good idea and access to the internet can search open materials databases, calculate the stability of candidate materials, predict new compounds, and even obtain a reasonable grasp on many of their properties. Such command of design is a powerful success story of MGI-inspired, DFT-based calculation engines such as the Materials Project [38], AFLOWLIB [39], and OQMD [40]. As but one example, searchable 'catalogs' of topological compounds [6] speak to the power with which theory and computation can democratize the virtual materials discovery process.

However, in many, if not most, cases these 'new materials' spend their lives *in silico*. Translating such discoveries to the laboratory and preparing them with needed quality and quantity for science and/or technology is a far more complex, time-consuming, and expensive proposition. One reason for this is because the corresponding 'synthesis by design' framework is largely unknown for inorganic solids, unlike organic molecules [41]. The promises and challenges to such a framework have been explored recently through workshops and their reports, such as the 2016 *Basic Research Needs Workshop on Synthesis Science* [42] sponsored by DOE Basic Energy Sciences and NSF's 2022 report *An Accelerated, Data-Driven, Materials Discovery Future* [43]. Key messages from such reports are that progress in both new synthetic tools (e.g., *in situ* probes of synthesis with feedback, access to extreme synthesis environments) and new synthesis concepts (e.g., non-classical crystallization, inorganic synthons) will be needed to move forward. Machines will aid and augment, but not replace, human ingenuity (what may we referred to above as 'smart assistants'). The pursuit of such synthesis research is best carried out in a balanced ecosystem comprised of small academic laboratories, midscale, focused centers, and large scale facilities. These sentiments were echoed by the MLRO workshop participants.

Materials synthesis in LoF was a significant part of the MLRO workshop discussion, and participants identified a wide-ranging set of questions that lie along the pathway to any synthesis by design framework. High priority issues in predictive synthesis emerged around areas of complexity (How do we synthesize multi-component solids with targeted properties? How do we design and grow multifunctional materials?), scale-bridging (How do we control compositional and structural heterogeneity at length scales beyond the atomic- and nanoscale?), defect control (How do we study defects in amorphous and disordered solids?), and synthesizability itself (How can we more deeply understand the factors that determine if a predicted compound can be made in desired form, and how?).

Behind these and other motivating research drivers lies a larger view that a big payoff in materials synthesis could come from a far more comprehensive understanding of kinetic control of synthesis and assembly, leading to the identification and isolation of nonequilibrium phases, compositions, and/or morphologies. A vast, unexplored space of functional materials, structures, and architectures is believed to lie in uncharted local minima of multi-dimensional energy landscapes. This *terra incognita* must be surveyed if we are to build better predictions for materials and how to synthesize them, e.g., through more complete databases. How can these energy landscapes be explored computationally and experimentally, minima identified, and kinetic pathways to them be fashioned in ways that lead to predictive synthetic outcomes? What role do extreme reaction environments play? How can automation, AI, and data-driven discovery take a leading role in this research? These questions offer both challenges to the LoT and a blueprint for the MLRO synthesis LoF.

Some answers are already emerging and being put into action (see sidebar "Crystal Growth: Deep Learning in the Zone"). However, for the most part, the MLRO synthesis LoT has not qualitatively changed in decades. Academic and government laboratories still depend largely on bespoke synthesis of research samples by students and postdocs, either serially or in small parallel batches. Incomplete natural language literature recipes, intuition, and trial-and-error often provide the synthetic strategy. There are notable exceptions, for example guided synthesis of Al_2O_3 surfaces using microcalorimetry as input to DFT simulations to establish the chemical environment [44] and the 'panoramic synthesis' of solid state materials using synchrotrons [45-48]. Both these examples underscore the importance of *in situ* and *operando* synthesis monitoring and are harbingers of the approaches that are certain to populate the MLRO LoF in predictive synthesis.

The present-day approach, while remarkably successful, reveals critical gaps in our understanding and our capabilities that, if overcome, would dramatically improve the synthesis hit rate and accelerate discovery. Workshop participants identified several prospective directions where the community could make progress. While a role for AI and ML-based autonomous discovery was considered a certainty, these were not considered the *sine qua non* for success.

- We lack effective ways to link predictions in simulation space to implementation in the laboratory. That is, how can we turn a simulated synthesis prediction into a real-world recipe? Conversely, experimental output should be more intelligently piped to simulations, allowing better decision-making (either by human or machine) for iterative simulation-synthesis loops.

- Our toolsets and algorithms for efficiently, rapidly, and economically probing high dimensional synthesis parameter spaces fall short of what is needed as this element of complexity becomes a key design criterion for new materials (e.g., multicomponent entropy stabilized materials). AI, robotics, and autonomation may offer a path forward through these complex spaces.

- Co-design of *in situ* and *operando* instrumentation at national facilities offers a powerful means to study transient atomic motions during growth and processing. These data can be used not only to direct synthesis at the beamline by correlating chemistry to control parameters, but also to provide input to mechanistic theories and models of the synthetic pathway itself. If verified, such mechanistic rules can apply beyond a single exemplar system, multiplying the impact on synthesis science and facilitating translation to the lab.

- Multimodal characterization of the synthesis process either in facilities or laboratories should capture results across multiple probes, populating reusable databases, feeding on-the-fly feedback, and potentially helping to steer an autonomous experiment itself. Such tools can provide means to streamline the critical identification of growth products for increasingly complex compositions and forms, particularly thin films and nanomaterials. Important will be finding ways to adapt in situ probes to the lab scale or identifying proxies that can be measured away from the facility.

- MLRO at the mesoscale often lie far from equilibrium, and guiding principles of assembly are lacking except in certain cases, for example MOFs. Fundamentally, our ability to study heterogeneous and transient mesoscale structures, especially in complex, multi-component systems, is underdeveloped and presents an enormous phase space of possible novel materials. Progress will require tools for quantitative measurements of defects and order/disorder on the mesoscale beyond imaging and tied to protocols for matching simulations to observed mesoscale order. Here the power of AI to extract features and to recognize hidden patterns may be a powerful ally in the LoF.

- Many materials have polymorphs that are very close energetically, but only one of the polymorphs is desired in a synthesis. Finding a route to nearly stable materials could occur by identifying links, or tipping points, in the process of synthetic transformations or nucleation and growth processes that favor the target while protecting against an undesirable thermodynamic sink. This naturally leads to a need to identify and steer critical synthetic events, presumably by direct observation and intervention, informed by simulation.

- Data provide critical input to synthesis design, whether that is by human or AI. Multiple gaps in our current data curation infrastructure stand in the way of the promise for either. For instance, we lack databases of synthesis protocols in a structured, machine-readable format, including inputs and outcomes (both success and failure); we lack searchable records of thin film syntheses on substrates, and we lack thermodynamic databases of metastable reaction intermediates.

3.4 Dynamics of atoms and electrons

Over recent decades it has become clear that driven systems offer unique opportunities to understand and leverage new phenomena, from nonequilibrium phase transitions to the flow of energy and entanglement among and between electronic and atomic degrees of freedom, to the transient situations that arise during crystal growth. To understand driven materials and materials far from equilibrium and to realize the potential they hold, critical advances are needed both experimentally and in theory and simulation.

The need for new experimental techniques to examine the dynamics of atoms and electrons in materials has been highlighted by the community, as in *Frontiers of Materials: A Decadal Survey (2019)* [49] by the National Academy of Sciences, and *Basic Energy Sciences Roundtable: Opportunities for Basic Research at the Frontiers of XFEL Ultrafast Science* [50]. Emerging methods, such as ultrafast and time-resolved x-ray and electron diffraction methods, give insights into the dynamics of lattice structure from the bulk down to nanoscale regions. This builds on decades of work that continues today using time-resolved optical probes that probe electronic degrees of freedom from picosecond to attosecond scales using spectroscopy and inference of the frequency-dependent optical conductivity.

With this progress, there remain significant gaps in capabilities. Many ultrafast methods are quite perturbative, involving significant deposition of energy into the material under study. This limitation makes it difficult to apply such approaches to some of the systems with the richest dynamics, and unusual emergent excitations such as spin liquids, systems with Majorana fermions, anyons, and so on, all of potential interest for quantum applications. Sensitive electronic measurements generally remain restricted to the GHz regime and below, far slower than the timescales of electron-electron (fs) and electron-phonon (ps) scattering within materials. Scanned probe techniques offer unprecedented spatial resolution at surfaces for characterizing structure, electronic spectroscopic (scanning tunneling microscopy), chemical information (atomic force microscopy), optical response (Scanning Near-field Optical Microscopy), local potentials (Superconducting Single-Electron Transistor), and magnetism (magnetic force microscopy, scanning SQUID microscopy and nitrogen vacancy centers). However, these techniques are usually slow and difficult to integrate with extreme environments of interest (high magnetic fields, high pressures, electrochemical environments).

Often it would be advantageous to have high spatial resolution real-time information about atomic positions *in situ* during materials growth. Traditional surface techniques such as Reflection high energy electron diffraction are helpful in epitaxial thin film growth but are slow and average over large areas, while faster and higher resolution methods are very difficult to combine with challenging growth environments. Often interfaces of critical interest are buried.

While cutting-edge techniques like four dimensional scanning electron microscopy (4D-STEM) [51] and scanning nanostructure electron microscopy (SNEM) [52] can provide great insights into structural and electronic degrees of freedom (e.g. through electron energy loss spectroscopy), sample preparation is lengthy, involved, and not well suited to characterization of structures during *in situ* and *operando* experiments.

Gaps in capabilities likewise remain in the modeling and computational dimension. Density functional theory and related methods have given tremendous insights into electronic structure and thermodynamic stability of materials. Applying these approaches across many scales and to disordered systems of mixed phases and constituents remains computationally expensive because of the need for huge supercells. Time-dependent techniques are usually similarly limited to timescales short compared to many experiments, and accounting for transient and excited states with accuracy remains a critical need. Extending the successes of computational approaches to equilibrium structure-function relationships to driven and nonequilibrium systems will be essential to realize the benefits of the LoF goals.

Bridging these gaps will be necessary to achieve the full potential of the LoF, with symbiotic reinforcement between computational modeling and laboratory implementation of materials synthesis, fabrication, characterization, and potential device development. Investment in the development of novel experimental and computational techniques that will cross the length and timescales inherent in materials growth and nonequilibrium response is essential.

4 Opportunities for Research, Training, and Access in LoF

Realizing the scientific goals laid out above will require a well-trained, diverse, and talented group of creative and productive researchers and technicians that pull from all segments of the population, including those that are historically underrepresented. Training in LoT involves close mentoring by laboratory PIs, and by mentoring of less experienced group members by more experienced ones. In the realm of experimental science, it resembles a hands-on apprenticeship beyond the more standard classroom training received by students, including problem and hypothesis definition, design of experiments, data analysis and knowledge extraction, manuscript preparation and so on. In this section we explore special opportunities in research, training and accessibility that were identified by MLRO participants related to LoF. We also discuss challenges inherent in realizing these opportunities.

4.1 Opportunities in recruitment, training and research offered by LoF, and challenges in getting there

Graduate students of tomorrow are more likely to be working in highly diverse *teams* of researchers from different disciplines, rather than as individuals pursuing a personal research goal. They must learn to interact effectively across disciplines and with people from very different training and experiential backgrounds. These developments present special challenges, but also great opportunities both in training in transferable skills [53] and in democratizing scientific discovery and lowering the barrier to participation.

Although in its infancy, remote access to equipment will be a major part of LoF. Researchers and students will be able to interact with and use equipment from afar through web-based tools that integrate to, and can in principle, operate hardware in the lab setting with AI assistance. One vision parallels the revolution in semiconductor chip manufacturing that accompanied the development of the foundry concept, in which external users could specify a chip design in software and the foundry could execute the steps to manufacture it. As we explore below, it opens up interesting possibilities in training and access. For want of a better description we will call this the "laboratory as a foundry" (LaaF) concept (another name that has been used is laboratories as a service, LaaS), though noting that LaaF will represent a small part of LoF in general.

With this said, some of the best experiences for students is actual hands-on work in a physical laboratory and this is likely to remain at the heart of LoF as well. In this regard, an interesting discussion centered about investing in research to find ways to make low-cost instrumentation. A key design or scientific innovation that allows an instrument to make a "good enough" measurement at an order of magnitude lower cost can be hugely impactful because it would allow the equipment to be deployed in an order of magnitude more labs across the country at the same cost. This approach to instrument development deserves attention. A combination of affordable access to equipment for smaller institutions and access to shared facilities can help a wider range of universities to gain access to research grade instrumentation, reaching many more communities, and could be especially impactful at reaching previously under-resourced communities.

4.2 Opportunities and Challenges in getting there

If equipment is bought not built, how do we pass on skills for innovating and building? Research is greatly accelerated when manufactured experimental equipment can be bought off the shelf. We can now work with computers without wiring up integrated circuits or even programming, which is fantastically enabling. We can buy rather than build dilution refrigerators, so people can do low-temperature measurements without being professional cryogenic engineers. However, this comes with costs. The experimental equipment becomes a black box. Many of the intricacies and approximations inherent in the design and manufacture are lost to the student using it. It becomes a challenge to give students a deep understanding of how best to use the equipment in their research and how to get the best out of it. This becomes even worse in the context of shared facilities. There may be opportunities for AI to mitigate the dangers of the black box by catching incorrect settings and mitigate at least for “known unknowns.” These AI guard-rails are currently largely absent and need research and investment. Providing incentives to manufacturers to invest in this could be valuable. However, beyond this, a correct balance of training in the background of “how things work,” with using machines as AI guarded black boxes, will be key.

The LoF presents special training opportunities in this regard, too. If a lab is operating as a LaaF, relatively low-cost practicals could be devised where students are actually solving problem sets that involve controlling real equipment rather than reading about it in books. These could be designed to include common misconceptions and mistakes in a way that allows the students to learn by actually making the mistakes and recovering from them. We also note that it is costly and time consuming for individual universities to develop advanced courses on research-grade methods. Understanding ways to do this effectively across institutions can be impactful, for example, national schools on advanced topics. These could possibly be organized by LaaF facilities with inputs from university-based experts. Such schools are already offered by several national user facilities, for example on synchrotron x-ray techniques [54]. New cooperation and funding models are needed to broaden this approach beyond schools organized by major facilities.

In a research setting, the LaaF model will free up student time for more creative thinking tasks which will accelerate discovery. It was pointed out that it will drive thinking in instrument design in the direction of greater modularity, which could increase collaboration prospects because of exchange of standardized samples and chambers between laboratories. To realize this model, community standards need to be developed. A challenge is that “One group’s sample is another group’s dirt.” Shared equipment needs to be carefully managed to avoid contamination. Modularity presents an opportunity in this regard.

A challenge with shared facilities is the slow and externally imposed access model (i.e., access is not driven by scientific expedients). Months or even years often pass between the idea for an experiment and its actual execution. This can decelerate research and dis-incentivize using those capabilities. Exploring different access models would result in accelerations in discovery.

Many of the issues of LaaF apply equally to data analysis and modeling, which is currently slow, repetitive and human intensive. Data analysis is often a rate limiting step in scientific discovery. Investments in the development, maintenance, and deployment of analysis software, including AI enabled, can be expected to result in large returns if they significantly shorten the time from measurement to knowledge extraction.

The challenge of integrating data capture with shared databases as part of the data analysis process is also hardly developed but will have significant benefits in terms of scientific reproducibility. It adds significant complication to the analysis software, but the potential gains in future AI enabled discovery are enormous. It will need up-front investments, but also a commitment, and financial and bureaucratic mechanisms, to continue to support the infrastructure for as long as it is useful. For example, infrastructure is needed for automated capture of metadata through QR-codes on instruments, and samples, and Internet of Things (IoT) connected instrumentation connected to shared database backends. A further opportunity is the use of AI for automating data capture. Recording speech, transcribing it to text, and beyond that to other representations (e.g., large language models like ChatGPT) provide a low overhead way for researchers to capture contextual information and metadata that can ultimately end up in a machine-readable format. These ideas can also be extended to video where both sound and images are interpreted. The development of augmented or virtual reality (AR,VR) and heads up display (HUD) technologies may also play a role in LoF.

Members of the workshop expressed major privacy and potential mental health concerns with developments such as these. They could lead to unwanted stress on researchers and risk avoidant behavior, which is undesirable since taking intellectual risks, safely, is an important part of scientific discovery. A no-blame culture would be essential because every mistake and misstep is recorded and attributable, in principle. These aspects of LoF need to be rolled out carefully, assessing the social, psychological, and legal aspects hand in hand with the technical ones.

It is difficult to assess what the future impact of large, shared databases of experimental data will be, in large part because we do not have them, and so demonstrations to date have not been transformational. It took half a century and multiple attempts for deep learning to show its abilities. Although many of the concepts and algorithms had been in place for decades, they didn’t result in transformative outcomes due to a lack of data and computing power, but took off once these became available. The capture and sharing of experimental scientific data is not at the point where we can expect transformative outcomes at present, and we won’t know how transformative they will be until we get there. Organization and investments are needed to get there.

There are inherently sociological, legal and financial issues to be solved when capturing community wide experimental data that will have to be addressed. What are the incentives for researchers to openly share hard-won data? Materials science is inherently different to astronomy, often cited as a model for open science, because of the huge intellectual, time and financial investments that individuals put into producing the samples and making bespoke measurements. What are the correct incentives to allow them to share the data in a way that will advance their careers? Who will pay to maintain the data stores? Who will pay (for example, cloud credits) to access it and maintain the infrastructure? There is a huge desire for data "paid for by taxpayers" to be open and available back to the taxpayer, but there is not a proper understanding of what this means, what it costs in time and money to make it open, nor what the benefit is to the taxpayers of having this access. This needs some focus. In general, investors want a return on their investment. In this case, the return to the investors, the taxpayers, need not necessarily be the data itself but the value that it produces. In a commercial setting the value is financial, but here there will also be non-fiduciary social benefits that need to be understood. It will be difficult to move forward at scale until these aspects are better understood, but the impacts of AI on scientific discovery will be limited until they are.

A robust FAIR data infrastructure will allow for new modalities for sharing information. The scientific literature has served us well for 350 years, with humans writing papers describing results in a form easily assimilated by other humans. However, the rate at which papers are appearing is exploding, doubling about every 15 years [55]. The scale of modern scientific literature makes it inherently unreadable by humans since the time we have for reading is not growing exponentially. This is a crisis since scientific papers are our primary vehicle for archiving and sharing scientific knowledge. Natural language processing models are being developed to allow machines to read papers written for humans. However, a goal for LoF must be that "papers" are written to be read by machines (as well as by humans) in the first place. Research is needed into how to do this effectively, both in terms of computational infrastructure, but also things as basic as knowledge representations and metadata schemas.

There are enormous potentialities, challenges and opportunities for access and democratization of science in LoF that are taking full advantage of modern AI/ML, automation and data analytic methods. Investments here can be expected to have large impacts, and can cut across all of the different scientific domains, but it will be important to manage the social, safety and political dimensions as this is rolled out.

5 Acknowledgements

We would like to thank the attendees and questionnaire respondents for willingly sharing their time, expertise, and vision. We would also like to thank the Division of Materials Research at the National Science Foundation for financial support and organizational help. Finally, we would like to thank Alexis Avedisian of the Data Science Institute at Columbia University for extensive technical, logistical, and administrative support.

6 References

[1] Christodoulou, J., Friedersdorf, L.E., Sapochak, L., Warren, J.A., 2021. The Second Decade of the Materials Genome Initiative. JOM 73, 3681–3683. https://doi.org/10.1007/s11837-021-05008-y

[2] Nikolaev, P., Hooper, D., Webber, F., Rao, R., Decker, K., Krein, M., Poleski, J., Barto, R., Maruyama, B., 2016. Autonomy in materials research: a case study in carbon nanotube growth. npj Comput Mater 2, 1–6. https://doi.org/10.1038/npjcompumats.2016.31

[3] https://nap.nationalacademies.org/catalog/12640/frontiers-in-crystalline-matter-from-discovery-to-technology

[4] https://dmref.org/

[5] Vergniory, M.G., Elcoro, L., Felser, C., Regnault, N., Bernevig, B.A., Wang, Z., 2019. A complete catalogue of high-quality topological materials. Nature 566, 480–485. https://doi.org/10.1038/s41586-019-0954-4

[6] https://www.topologicalquantumchemistry.com/#/

[7] Zhang, T., Jiang, Y., Song, Z., Huang, H., He, Y., Fang, Z., Weng, H., Fang, C., 2019. Catalogue of topological electronic materials. Nature 566, 475–479. https://doi.org/10.1038/s41586-019-0944-6

[8] Nanotechnology Research Directions for Societal Needs in 2020: Retrospective and Outlook (Science Policy Reports), Edited by Mihail Roco, Chad Mirkin and Mark Hersam, Springer, 2011.

[9] Nanofabrication: principles, capabilities and limits", Cui Zheng, Springer, Germany (2016).

[10] Midolo, L., Schliesser, A. and Fiore, A., 2018. “Nano-opto-electro-mechanical systems.” Nature nanotechnology, 13(1), p.11.

[11] Rodriguez, A.W., Hui, P.C., Woolf, D.P., Johnson, S.G., Lončar, M. and Capasso, F., 2015. “Classical and fluctuation-induced electromagnetic interactions in micron-scale systems: designer bonding, antibonding, and Casimir forces.” Annalen der Physik, 527(1-2), pp.45- 80.

[12] Maragò, O.M., Jones, P.H., Gucciardi, P.G., Volpe, G. and Ferrari, A.C., 2013. “Optical trapping and manipulation of nanostructures.” Nature nanotechnology, 8(11), p.807; Samorì, P. ed., 2006. Scanning probe microscopies beyond imaging: manipulation of molecules and nanostructures. John Wiley & Sons.

[13] https://science.osti.gov/bes/Community-Resources/Reports

[14] A. M. J. Zwerver et al, Qubits made by advanced semiconductor manufacturing, Nature Electronics (2022). DOI: 10.1038/s41928-022-00727-9

[15] Awschalom, David, et al. “Development of quantum interconnects (quics) for next-generation information technologies.” PRX Quantum 2.1 (2021): 017002.

[16] M. Bowick, A. Kosmrlj, D. Nelson, R. Sknepkek, “Non-Hookean statistical mechanics of clamped graphene ribbons”, Phys. Rev. B 95, 104109 (2017).

[17] Košmrlj, A., Nelson, D.R., 2016. “Response of thermalized ribbons to pulling and bending.” Phys. Rev. B 93, 125431 (2016).
[18] Chakraborty, Suman, ed. ‘Mechanics over micro and nano scales.” Springer Science & Business Media, 2011.
[19] Blees, M.K., Barnard, A.W., Rose, P.A., Roberts, S.P., McGill, K.L., Huang, P.Y., Ruyack, A.R., Kevek, J.W., Kobrin, B., Muller, D.A., McEuen, P.L. “Graphene kirigami.” Nature 524, 204–207 (2015).
[20] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter. “Molecular biology of the cell.” Garland Science, New York, 6th edition, 2015.
[21] W. S. Wong and A. Sallero, editors. “Flexible Electronics: Materials and Applications.” Springer, New York, 2008.; S. Xu, Z. Yan, K.-I. Jang, W. Huang, H. Fu, J. Kim, Z. Wei, M. Flavin, J. McCracken, R. Wang, A. Badea, Y. Liu, D. Xiao, G. Zhou, J. Lee, H. U. Chung, H. Cheng, W. Ren, A. Banks, X. Li, U. Paik, R. G. Nuzzo, Y. Huang, Y. Zhang, and J.A. Rogers. “Assembly of micro/nanomaterials into complex, three-dimensional architectures by compressive buckling”. Science, 347:154–159, 2015.
[22] Andre K., and Konstantin S. Novoselov. “The rise of graphene.” Nature materials 6.3 (2007): 183-191.
[23] Annesini, Maria Cristina, et al. Artificial organ engineering. Berlin/Heidelberg, Germany: Springer, 2017.
[24] Surjadi, James Utama, et al. “Mechanical metamaterials and their engineering applications.” Advanced Engineering Materials 21.3 (2019): 1800864.
[25] Reid, MT Homer, Alejandro W. Rodriguez, and Steven G. Johnson. “Fluctuation-induced phenomena in nanoscale systems: harnessing the power of noise.” Proceedings of the IEEE 101.2 (2012): 531-545; French, Roger H., et al. “Long range interactions in nanoscale science.” Reviews of Modern Physics 82.2 (2010): 1887.
[26] Decca, Ricardo, Vladimir Aksyuk, and Daniel López. “Casimir Force in Micro and Nano Electro Mechanical Systems”, Casimir Physics. Springer Berlin Heidelberg, 2011; Ricardo Decca and Daniel López.”
[27] Nguyen, Clark T-C. “MEMS technology for timing and frequency control.” IEEE transactions on ultrasonics, ferroelectrics, and frequency control 54.2 (2007): 251-270.
[28] Chaste, J., et al. “A nanomechanical mass sensor with yoctogram resolution”, Nature Nanotechnology 7 (2012): 301-304.
[29] Rugar, Daniel, et al. “Single spin detection by magnetic resonance force microscopy”, Nature 430 (2004): 329-332.
[30] Antonio, Dario, Damián H. Zanette, and Daniel López. “Frequency stabilization in nonlinear micromechanical oscillators.” Nature communications 3.1 (2012): 806
[31] Barnard, Arthur W., et al. “Real-time vibrations of a carbon nanotube.” Nature 566.7742 (2019): 89-93.

[32] Imboden, Matthias, and David Bishop. “Nanomanufacturing.” Physics Today 67.12 (2014): 45; Lally, Richard, et al. “The Construction of a Fully Integrated, MEMS Based, Atomic-Scale 3D Printer”
[33] Malshe, A. P., et al. “Tip-based nanomanufacturing by electrical, chemical, mechanical and thermal processes.” CIRP annals 59.2 (2010): 628-651
[34] Hiramoto, Toshiro. “Five nanometre CMOS technology.” Nature Electronics 2.12 (2019): 557-558
[35] Zwerver, A.M.J., Krähenmann, T., Watson, T.F. et al. “Qubits made by advanced semiconductor manufacturing.” Nature Electron 5, 184–190 (2022).
[36] D. Wan, D. R. Nelson and M. J. Bowick, “Thermal stiffening of clamped elastic ribbons”, Physical Review B96, 0114106 (2017).
[37] https://www.nist.gov/news-events/news/2022/09/nist-and-google-create-new-supply-chips-researchers-and-tech-startups
[38] https://materialsproject.org
[39] https://www.aflowlib.org
[40] https://oqmd.org
[41] Segler, M.; Preuss, M.; Nature, M. W. Planning Chemical Syntheses with Deep Neural Networks and Symbolic AI. Nature 2018, 555 (7698), 604 610. https://doi.org/10.1038/nature25978.
[42] https://www.osti.gov/servlets/purl/1 616513
[43] https://www.materialsarchive.org/reports/2DILM_Report_Final_Public.pdf
[44] A.R. Bielinski, E.P. Kamphaus, L. Cheng, and A.B.F. Martinson, “Resolving the Heat of Trimethylaluminum and Water Atomic Layer Deposition Half-Reactions,” J. Am. Chem. Soc. 144, 15203 (2022).
[45] D.P. Shoemaker, D.Y. Chung, J.F. Mitchell, T.H. Bray, L. Soderholm, P.J. Chupas, and M.G. Kanatzidis, “Understanding Fluxes as Media for Directed Synthesis: In Situ Local Structure of Molten Potassium Polysulfides,.” J Am. Chem. Soc. 134, 9456 9463 (2012).
[46] Shoemaker, D. P.; Hu, Y.-J.; Chung, D. Y.; Halder, G. J.; Chupas, P. J.; Soderholm, L.; Mitchell, J. F.; Kanatzidis, M. G. In Situ Studies of a Platform for Metastable Inorganic Crystal Growth and Materials Discovery. Proc National Acad Sci 2014, 111 (30), 10922–10927. https://doi.org/10.1073/pnas.1406211111.
[47] Vasquez, G.; Huq, A.; Latturner, S. E. In Situ Neutron Diffraction Studies of the Metal Flux Growth of Ba/Yb/Mg/Si Intermetallics. Inorg Chem 2019, 58 (12), 8111–8119. https://doi.org/10.1021/acs.inorgchem.9b00857.
[48] Franco, C.; Wustrow, A.; Xia, B.; Baccarella, A. M.; Burgos, F.; Nicasio, J.; Dooryhee, E.; Neilson, J. R.; Simonson, J. W. Optimized in Situ Crystal Growth and Disordered Quasi-One-Dimensional Magnetism in Li2Mn2(MoO4)3. Phys Rev Mater 2020, 4 (4), 045404.
[49] https://www.nationalacademies.org/our-work/frontiers-of-materials-research-a-decadal-survey
[50] https://www.osti.gov/biblio/1616251
[51] Ophus, C., 2019. Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond. Microscopy and Microanalysis 25, 563–582. https://doi.org/10.1017/S1431927619000497

[52] Rakita, Y., Hart, J.L., Das, P.P., Shahrezaei, S., Foley, D.L., Mathaudhu, S.N., Nicolopoulos, S., Taheri, M.L., Billinge, S.J.L., 2023. Mapping structural heterogeneity at the nanoscale with scanning nano-structure electron microscopy (SNEM). Acta Materialia 242, 118426. https://doi.org/10.1016/j.actamat.2022.118426
[53] https://www.lifescied.org/doi/10.1187/cbe.20-12-0276
[54] https://neutrons.ornl.gov/nxs
[55] Bornmann, L., Haunschild, R., Mutz, R., 2021. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. Humanit Soc Sci Commun 8, 1–15. https://doi.org/10.1057/s41599-021-00903-w

7 Appendices

7.1 Organizational details

1. The workshop took place on the 16-17th November 2022 virtually
2. The meeting consisted of one invited lecture segment and 5 breakout sessions.
3. Input was sought from the community through a questionnaire
    a. Questionnaire was advertised to the community by
        i. NSF-DMR mailing to subscribers
        ii. Materials science department chairs mailing list
        iii. Chemistry department chairs mailing list
        iv. Physics department chairs mailing list
        v. MRS - A link to the community LinkedIn
        vi. APS - DCMP, DMP, GMAG mailing lists
        vii. Plus additional mailing lists such as the conference attendees of the 2022 Machine Learning in Science and Engineering conference data science and dean mailing lists.
4. There were 52 meeting attendees by invitation only. Invitees were picked from the questionnaire respondents and from lists of names suggested the organizers, advisory committee members and other stakeholders. Final decisions on invitations were then made based to maintain diversity in scientific expertise, gender, geography, institution type, and to ensure representation from under represented minorities and members of mid and large-scale experimental facilities.

5. Discussion topics for the meeting were harvested from the questionnaire responses, and from the invited attendees, and were presented in the breakouts in the form of powerpoint discussion prompts

6. Overview of meeting agenda and organization

7. This report represents the results of the questionnaire and the discussions at the workshop

7.2 Questionnaire

Introduction:

The condensed matter and materials science communities stand at a time of rapid changes in how materials can be made, how experiments are run, and how data are collected, handled, and processed. In preparation for an NSF-sponsored workshop on "**Materials with Long Range Order**", part of a series on **"Materials Laboratories of the Future: Instrumentation and Infrastructure to Accelerate the Unification of the Materials Innovation Infrastructure"**, we seek input from the condensed matter and materials communities to help us as organizers understand how these new opportunities in materials research should be identified, resourced, developed, deployed, and coordinated. We want to know what *your* vision of "Materials Labs of the Future" is. What instrumentation, what capabilities, what new modalities of interacting with data do you see as important to you as an individual researcher and to the community at large?

You should not not feel constrained by the zeitgeist of AI, big data, etc. (though these may indeed be critical to your vision). Rather, explore through the questions below what you see as the path(s) to advance the fields of condensed matter and materials science research over the next decade and beyond.

We invite you to answer as many or as few of the questions below as you wish with whatever level of detail you feel appropriate. Your answers will be a valuable input to the discussions at the workshop and to the report that emerges.

We the workshop organizers appreciate the time that it takes to give thoughtful responses to questions such as these. Thank you in advance for your time and for your willingness to contribute your insights.

Simon Billinge

John Mitchell

Susanne Stemmer

Questions:

What do you think are the most important scientific goals, questions and challenges to address in condensed matter and materials sciences in the next 5-10 years?

What tools, capabilities, and expertise are needed to meet these goals, and of those what is now lacking?

How would you envisage a lab of the future, and how can it help to meet the scientific goals? What are the most significant resources, infrastructure capabilities, and investments needed to realize this vision. If you posed several goals, and if you feel they call for different lab structures / approaches, feel free to respond separately for each goal, here and in questions below: “Goal 1: Lab structure … ; Goal 2: Lab structure …”

What role do you see Artificial Intelligence/Machine Learning (AI/ML), Big Data, automation and autonomation playing in the labs of the future? Do you have examples where these are already important enablers of research progress?

Do you anticipate or have you already observed limitations, hazards , and downsides associated with the use of these approaches? Include not only physical/health hazards but also scientific hazards such as biases/blindspots.

What other developments do you see as being important to the labs of the future?

What partnerships of teams and expertise (for example, experimental, computational, engineering, industry involvement) will be needed to realize your vision for the labs of the future? How may these be organized to increase the materials community’s access to the capabilities (e.g. capabilities shared across several research groups, physical and/or virtual centers, user facilities, computer scientists and engineers funded to build open source tools as a service, other?)

Have any recent developments (technical, organizational, etc.) been underwhelming or not fully convincing to you so far with regards to their impact? Do you have questions that could help proponents articulate impact, or suggestions that could increase impact?

What novel funding modalities are needed to realize the labs of the future?

What novel training will participants in the labs of the future need and how will they get it?

Are there special opportunities from your vision for a lab of the future that could broaden participation (increase participation by under-represented groups) and how can these opportunities best be realized?

What other thoughts not specifically covered by the questions above would you like to share with us?

7.3 Meeting Agenda

November 16 - 17, 2022 - VERSION 10 (V10)

Pre-Session 1 will include: Welcome from organizers; a message from L. Sapochak (DMR); "Automation, Autonomation, and Autonomy" from B. Maruyama; and Q&A

Session 1 (Wednesday, November 16)

Breakout Deliverable:

1. A bullet list of science questions/problems in MLRO that cannot be answered now due to a lack of experimental capability, or supporting theory/modeling.

2. For each question, a sub-list of specific exemplars.

3. For each exemplar, a list of the reasons it cannot be solved currently.

- Example: **top level**: pick a class of a class of science/engineering problems (e.g., "topological qubits" or "Interfacial superconductivity", "spin liquids", "lower-power electronics"), or instrumentation or experimental approaches (e.g., "thin film growth" or "low-temperature measurements"). **Next level** pick concrete exemplar problems, such as "experimental signatures of Majorana quasiparticles in hybrid structures", "large area lattice-matched substrates for growth of topological material x", or "quantify structural relaxations at heterointerfaces in three dimensions", "ferroelectric FET for low power electronics".

    - For each exemplar, identify why the problem cannot be solved currently (the capability gap(s)). Remember, focus the discussion on why the question CANNOT be answered currently and NOT YET on how a LOF can be built to answer it.

Session 2 (Wednesday, November 16)

Breakout Deliverable:

Each breakout group is assigned exemplars from session 1. Present a list of technical capabilities that would be needed to solve these examples. Capture discussion surrounding the decisions as much as possible.

- **If a solution involving robotics is chosen:** (a) discuss if the robotics are transformational (the solves a problem that couldn't be solved without them) or a convenience/efficiency. (b) Identify which one of the following three cases the robots would provide: (i) just automation (increased throughput, greater precision and reproducibility), (ii) an autonomated automation (a machine controlled optimization problem that makes the high throughput search converge in fewer steps), (iii) autonomous discovery (machine controlled search for unknown solutions/hypothesis testing).

- **If a solution not involving robotics is chosen:** list the improved experimental capabilities (hardware, algorithms, software, database infrastructure, integrations, etc.) that would have to be in place for success and any changes in experimental workflow/design.

<u>Session 2 Report Back:</u> We aim to create a list of novel capabilities that have been identified.

Session 3 (Wednesday, November 16) | Breakout Deliverable

Breakout Deliverable:
Each group will be given a list of novel capabilities associated with science exemplars, generated in the prior sessions. Deliver a bullet list strategic plan for realizing these capabilities as a lab of the future (LOF) by listing answers to each of these questions (and any other questions that come up in the discussions):

What does success look like? Keeping that in mind, answer these other questions below

1. How would the LOF be organized? Centralized? Decentralized? Distributed? Generically, where would it be located (regional, national)?

2. What would it cost (magnitude/scale)? How will it be paid for? How could costs be reduced?

3. How would it be accessed by a broad community?

4. How would it be maintained?

5. How would it be developed/extended?

What role would be played by existing NSF and DOE funded user facilities?
Session 3 Report Back: Using these thoughts, we aim to create a list of strategic visions.

Session 4 (Thursday, November 17)

Breakout Deliverable:

Answer to the question "What does your lab of the future look like?"

- Unleash your imagination, let yourself off the leash, think big,
- Deliver a vision that is so exciting that Congress will throw money at it, NSF just has to ask
- What will your grad student's life be like? Will they be working alongside robots, or working from home but running experiments remotely? Will they spend more time at different facilities away from home?
- What will your lab look like? Will it be the same but more shiny? Or radically different?
- How will your grad students interact with each other? Collaborators at other institutions? With You? With the machines in the lab? Will the lab notebook be a video that captures everything and AI that interprets it?
- What else excites you about labs of the future? What have we missed?
- But most importantly, how will all this allow you to deliver the science questions identified in Session 1 (or later)

Session 5 (Thursday, November 17) | Breakout Deliverable

Breakout Deliverable:

We will have an open discussion in the breakouts. Attendees may raise any top of mind issues. Please bring any new ideas that bubbled up overnight (new exemplars or tweaks to capabilities and strategic visions). We encourage a deeper discussion on cost and value propositions.

Session 5 Report Back: Session leaders will review and share any new thoughts from the groups. We aim to have a discussion on what an executive summary will look like.

Workshop Co-Chairs, Organizing Committee, and Session Scribes will attend a private, brief organizational meeting (20 minutes) to discuss compiling the post-workshop report. Please find the meeting details in your calendar and email.

7.4 Attendees

Charles	Ahn	Yale
Mohsen	Asle Zaeem	NSF
Luis	Balicas	FSU
Nazanin	Bassiri-Gharb	NSF
Simon	Billinge	Columbia
Cossima	Boswell-Koller	NSF
Mun	Chan	NSF
Karena	Chapman	Stony Brook
Vincent	Crespi	PSU
Tomasz	Durakiewicz	NSF
Lauren	Garten	Georgia Tech
David	Goldhaber-Gordon	Stanford
Rachel	Goldman	U Michigan
Alexander	Grutter	NIST
Tessema	Guebre	NSF
Julia	Hsu	UT Dallas
Pinshane	Huang	UIUC

Germano	Iannacchione	NSF/NAF/ MPS/DMR
Eugenia	Kharlampieva	NSF
Divine	Kumah	NC State
Dhananjay	Kumar	NCA&T
Paul	Lane	NSF
Jeanie	Lau	OSU
Stephanie	Law	PSU
Mingda	Li	MIT
Daniel	Lopez	Penn State
Paul	Maggard	NCSU
Mike	Manfra	Purdue
Andrew	Mannix	Stanford
Benji	Maruyama	Air Force RL
Steve	May	Drexel
Tyrel	McQueen	Johns Hopkins
Delia	Milliron	UT Austin
Austin	Minnich	Caltech
John	Mitchell	Argonne
Joseph	Montoya	Toyota Research Inst
Doug	Natelson	Rice University
Robert	Nemanich	ASU
Kate	Page	U Tenessee
Siddharth	Rajan	OSU

Joan	Redwing	PSU
Trevor	Rhone	Harvard
John	Schlueter	NSF
Leslie	Schoop	Princeton
Daniel	Shoemaker	UIUC
Susanne	Stemmer	UCSB
Sandhya	Susarla	ASU
Eric	Toberer	Colorado School of Mines
Stephen	Wilson	UCSB
Keisuke	Yazawa	Colo. School of Mines
Charles	Ying	NSF
Hongping	Zhao	OSU

7.5 Virtual Workshop photographs


Simon Billinge
THE DATA SCIENCE INSTITUTE AT COLUMBIA
Alexis Avedisian
John Mitchell
Steve May
Robert Nemanich
Daniel Lopez
Rachel S Goldman
Jeanie Lau
Vincent Henry Crespi
Divine Kumah
Kei Yazawa
Douglas Natelson
Sandhya Susarla
Julia Hsu
Luis Balicas
Katharine Page
Dhananjay Kumar
Eugenia Kharlampieva
Mohsen Asle Zaeem
Stephanie Law
Leslie Schoop
Susanne Stemmer
Cosima Boswell-Koller
Delia Milliron (she/her)
Austin J. Minnich


Douglas Natelson
Sandhya Susarla
Julia Hsu
Luis Balicas
Katharine Page
Dhananjay Kumar
Eugenia Kharlampieva
Mohsen Asle Zaeem
Stephanie Law
Leslie Schoop
Susanne Stemmer
Cosima Boswell-Koller
Delia Milliron (she/her)
Austin J. Minnich
Hongping Zhao
Pinshane Huang
Mingda Li
Souleymane Diallo (NSF)
Alex Grutter
andy
STRONGER TOGETHER
Paul Lane (NSF)
John Schlueter
Siddharth Rajan (Ohio S...
David Goldhaber-Gordon
Connecting to audio