Highlights from the Nanotechnology Knowledge Infrastructure (NKI) NSI

March 2017

Nanotechnology has the potential to help solve global challenges by generating and applying new multidisciplinary knowledge of nanoscale phenomena and engineered nanotechnology-enabled materials, structures, and products. Data underlying this new knowledge are vast, disconnected, and challenging to integrate into a broad scientific body of knowledge. The purpose of the NKI Nanotechnology Signature Initiative (NSI) is to leverage and extend existing and emerging resources, programs, and technologies to support the broader goals of the National Nanotechnology Initiative (NNI) by creating an infrastructure to accelerate the vetting of new knowledge and to enable effective data utilization. The goal of the NKI NSI is to coordinate the nanoscale science, engineering, and technology communities around the fundamental, interconnected elements of collaborative modeling, a cyber toolbox, and data infrastructure that will capitalize on American strengths in innovation, shorten the time from research to new product development, and maintain U.S. leadership in the sustainable design of engineered nanomaterials.

The NKI NSI was launched in May 2012 to provide a community-based, solutions-oriented knowledge infrastructure to accelerate nanotechnology discovery and innovation. The NKI white paper[1] describes four thrust areas that focus the efforts of the 11 participating agencies[2] on cooperative, interdependent R&D:

1. Build a diverse, collaborative community of scientists, engineers, and technical staff to support research, development, and applications of nanotechnology to meet national challenges.
2. Foster an agile modeling network for multidisciplinary intellectual collaboration that effectively couples experimental basic research, modeling, and applications development.
3. Build a sustainable cyber toolbox to enable effective application of models and knowledge to nanomaterials design.
4. Create a robust digital nanotechnology data and information infrastructure to support effective data sharing, collaboration, and innovation across disciplines and applications.

Nanoinformatics is inherently a community-based topic. As such, the agencies participating in the NKI NSI and non-Federal groups have undertaken a wide range of activities to address the needs of the nanotechnology research and development (R&D) community. The following paragraphs provide select examples of advancements that have been made toward the objectives of the NKI NSI since its launch, including key intersections with the broader R&D community.

Agencies participating in the NKI NSI and the National Nanotechnology Coordination Office (NNCO) foster the development of a **diverse, collaborative nanoinformatics community** by monitoring and engaging with other relevant efforts, supporting mechanisms and events that promote community dialogue, performing outreach, and educating a collaborative workforce.

[1] www.nano.gov/NSINKI

[2] Consumer Product Safety Commission, Department of Defense, Department of Energy, Environmental Protection Agency, Food and Drug Administration, National Aeronautics and Space Administration, National Institutes of Health, National Institute for Occupational Safety and Health, National Institute of Standards and Technology, National Science Foundation, and Occupational Safety and Health Administration.

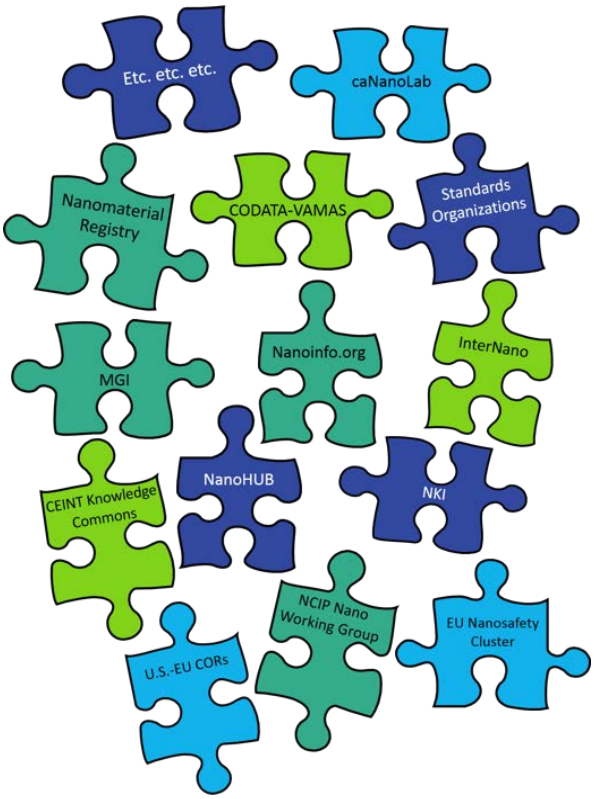

Figure 1. A wide variety of groups contribute to the nanoinformatics landscape.

The task of monitoring and engaging with other relevant efforts is complicated by the fact that the nanoinformatics landscape is extensive and constantly evolving (Figure 1). Some efforts include nanoinformatics as part of a larger nanotechnology-focused activity. For example, one of the seven U.S.-EU Communities of Research[3] addresses Databases and Computational Modeling for NanoEHS (nanotechnology-related environmental, health, and safety). Other activities, such as the Materials Genome Initiative (MGI),[4] are focused on materials informatics more broadly but are grappling with many of the same issues as the nanoinformatics community. NKI NSI participants share information and engage with these related activities and efforts where appropriate. For instance, several representatives of agencies participating in the NKI NSI also contribute to the Nanotechnology Working Group,[5] an interest group supported by the National Cancer Institute (NCI) National Cancer Informatics Program (NCIP). Members of the Nanotechnology Working Group have collaboratively written a series of consensus papers focused on nanoscale material and system data as part of a Nanomaterial Data Curation Initiative.[6] These papers provide a snapshot of current curation practices and concerns aimed at encouraging the widespread adoption of standardized practices.

NKI participants also frequently share information with the broader nanotechnology community. This outreach includes incorporating NKI NSI themes and activities into invited presentations at conferences and meetings with grantees, as well as hosting sessions and town hall discussions at relevant meetings and conferences to raise awareness of the NKI Signature Initiative. As described below, these conversations led directly to a productive collaboration between the Nanomaterial Registry[7] and nanoHUB.[8]

Modeling is the essential mechanism that couples the conceptual understanding of nanomaterials with the empirical knowledge gained from experimentation. Agencies participating in the NKI NSI contribute to the development of an **agile modeling network** designed to facilitate commercialization of nanotechnology discoveries while minimizing risks to humans and the environment.

[3] www.us-eu.org
[4] www.mgi.gov
[5] nciphub.org/groups/nanowg
[6] nciphub.org/groups/nanotechnologydatacurationinterestgroup
[7] nanomaterialregistrystage.rti.org/
[8] nanohub.org/

The National Science Foundation (NSF) has supported the creation of several key modeling resources, including the Knowledgebase of Interatomic Models (OpenKIM)[9] and nanoHUB. OpenKIM is an online suite of open source tools for molecular simulation of materials. This community resource supports various activities to train those doing simulations and to advance the use of molecular dynamics simulation to gain understanding and contribute to the design of materials. nanoHUB provides a virtual platform for computational nanotechnology research, education, and collaboration, and hosts over 320 simulation tools for nanoscale phenomena that run in the cloud and are accessed through a web browser. Agencies participating in the NKI NSI continue to add to the ongoing evolution of nanoHUB. For example, based on vibrant discussions at an NKI NSI-hosted town hall in 2013, a Nanomaterial Registry Portal[10] was added on nanoHUB (see page 5 for more information), enabling scientists to seamlessly access, analyze, and model the well-curated Nanomaterial Registry data using nanoHUB tools and resources. Only a small fraction of chemicals in commercial use have been evaluated fully for potential human health effects because current chemical testing is expensive and time consuming. Through its computational toxicology (CompTox) research, the Environmental Protection Agency (EPA) is working to predict potential toxicity of chemicals and to develop a cost-effective approach for prioritizing the thousands of chemicals that need toxicity testing. A major part of EPA's CompTox research is the Toxicity Forecaster (ToxCast),[11] which has evaluated over 2,000 chemicals from a broad range of sources.

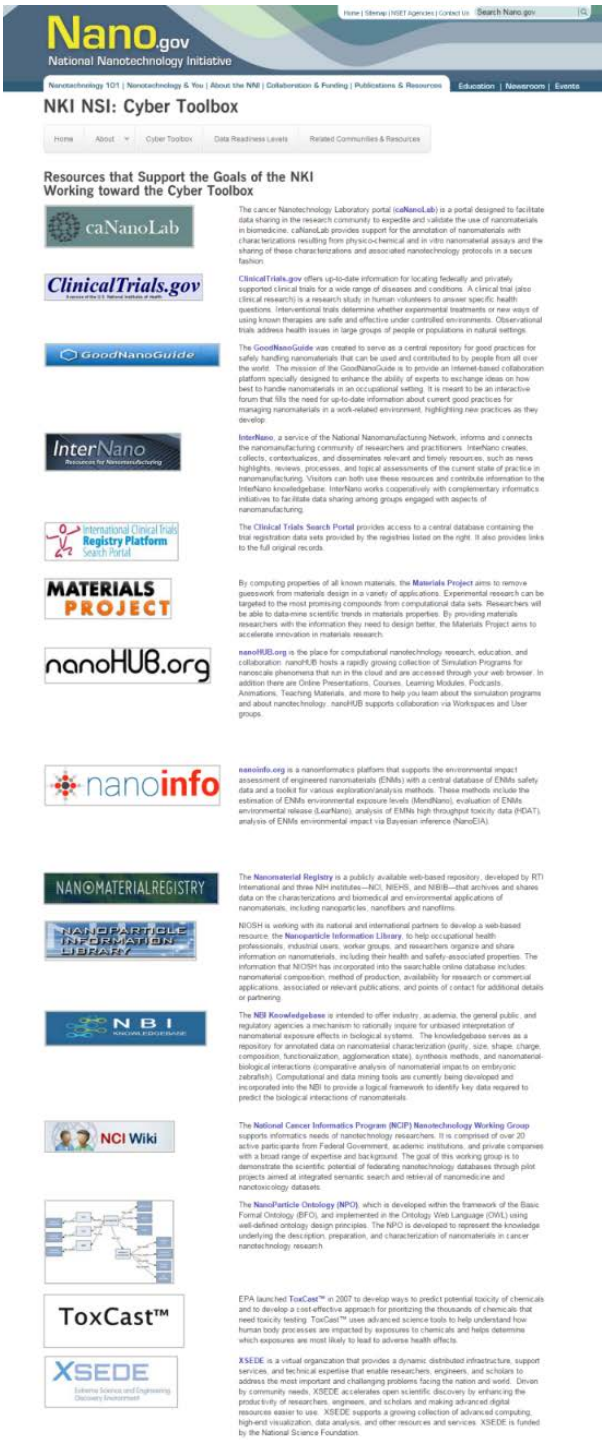

Figure 2. Resources on nano.gov that support the cyber toolbox and data and information infrastructure.

The development of a **sustainable cyber toolbox** is essential to the success of the NKI Signature Initiative. In support of this thrust, NNCO, with input from the agencies collaborating in the NKI NSI, has created and regularly updates a webpage (Figure 2) with information on more than a dozen

[9] openkim.org/
[10] nanohub.org/groups/nanomaterialregistry
[11] www.epa.gov/chemical-research/toxicity-forecasting

computational tools and other relevant resources that facilitate the analysis of experiments and the understanding of nanomaterials.[12] Several modeling tools (discussed above) and data and information infrastructure tools (discussed below) also include cyber toolbox elements.

NSF, through the Software Infrastructure for Sustained Innovation (SI^2) solicitation,[13] has supported and continues to support software for the simulation of nanosystems. Software development supported under SI^2 includes new software development, as well as enhancements to existing tools of the nanotechnology community, including VASP, Quantum Espresso, AMBER, and GAMESS. Researchers at the Army Engineer Research and Development Center (ERDC) developed a tool,[14] NanoGRID (Nano Guidance for Risk Informed Deployment),[15] that uses a five-tiered process to evaluate the potential hazard of and exposure risk associated with a nanotechnology-enabled product or process. NanoGRID has the potential to increase the capabilities of the Department of Defense and industry to reduce the time and cost of acquiring these technologies. The Materials Project,[16] which also supports the goals of the MGI, provides open, web-based access to computed information on known and predicted materials, including more than 500,000 nanoporous materials, as well as powerful analysis tools to innovate and design novel materials. Initiated in October 2011 as a joint collaboration between the Massachusetts Institute of Technology and the Department of Energy's Lawrence Berkeley National Laboratory, the Materials Project aims to remove the guesswork from materials design in a variety of applications and now has more than 20,000 users.

The nanotechnology community benefits from the ongoing development of a robust, integrated **nanotechnology data and information infrastructure**. A key component of this infrastructure is the development of a framework to share vetted nanotechnology data and information via standardized formats and vocabulary, to provide data access to the broader research community, and to augment and accelerate scientific discovery and innovation.

A critical component of sharing data among researchers and technology developers is a common nomenclature for describing and understanding the maturity and quality of the data. Toward this goal, representatives from the collaborating agencies developed a Data Readiness Level[17] (DRL) discussion document in 2013. Analogous to Technology Readiness Levels, the DRLs provide a shorthand method for conveying coarse assessments of data maturity from experiments or model predictions for use in improving analytical methods and validating or calibrating models, and for comparisons with legacy datasets.

The National Institutes of Health (NIH) have also conducted focused outreach to scientific publishers that has led to discussions among multiple publishing companies—including Nature Publishing Group, PLOS journals, the American Chemical Society Publications, and Elsevier—of possible implementation of shared policies for the deposition of nanoparticle characterization data into databases. For example, *Scientific Data,* Nature Publishing Group's journal of data descriptors, mandates the release of datasets that

[12] www.nano.gov/NKIPortal/CyberToolbox
[13] www.nsf.gov/funding/pgm_summ.jsp?pims_id=503489
[14] This tool currently is available for internal ERCD use only.
[15] Z. A. Collier *et al.*, Tiered guidance for risk-informed environmental health and safety testing of nanotechnologies, *Journal of Nanoparticle Research* **17**(3), 1–21 (2015).
[16] www.materialsproject.org
[17] www.nano.gov/NKIPortal/DRLs

accompany manuscripts[18] into approved data repositories, including caNanoLab.[19] Further, the scientific publisher Elsevier has implemented reciprocal linking between datasets housed in caNanoLab and the associated published research article.[20] This linkage is in place for ten journals with plans for additional journal linkages.

Agencies participating in the NKI NSI have supported the establishment of key components of the nanotechnology data and information infrastructure. The National Institute for Occupational Safety and Health supports the GoodNanoGuide[21] on nanoHUB. The goal of the GoodNanoGuide is to create a central repository for good practices for safely handling nanomaterials and to provide a collaboration platform where experts can exchange ideas on how best to handle nanomaterials in occupational settings. NIH has directly developed and supported the creation of two key databases: the NCI cancer Nanotechnology Laboratory portal (caNanoLab) and the Nanomaterial Registry. caNanoLab is a portal designed to facilitate information sharing to expedite and validate the use of functional nanomaterials in biomedicine. The Nanomaterial Registry provides access to curated information on the interactions of nanomaterials with biological and environmental systems, including links to related publications, modeling tools, computational results, and manufacturing guidance. The registry draws inputs from existing curated databases, including caNanoLab, and currently includes over 2,000 particle entries. In support of the MGI, NIST is developing and maintaining the Materials Data Repository,[22] the Materials Data Curation System,[23] and the Materials Resource Registry.[24] These tools and derivative capabilities can be leveraged by the NKI NSI. In particular, the Materials Resource Registry is being developed to provide end users with a search capability that enables access to materials data repositories around the world. The framework will be able to accommodate existing repositories as well as data from individual researchers and research organizations.

The NKI NSI leverages and extends existing and emerging resources, programs, and technologies to support the broader goals of the NNI. Significant progress has been made towards these goals by laying the foundation for an evolving infrastructure to accelerate the vetting of new knowledge and to enable effective data utilization. The agencies participating in the NKI NSI have fostered the development of a robust nanoinformatics community and have engaged with this community by organizing and participating in a wide range of relevant events, including a webinar series in the fall of 2015.[25] Several representatives of agencies participating in the NKI NSI and NNCO staff members also participate in the Nanotechnology Working Group and the U.S.-EU nanoEHS Communities of Research. NKI NSI participants and NNCO staff have also served on the steering committee for the three annual Nanoinformatics Workshops[26] that have

[18] www.nature.com/sdata/policies/repositories
[19] cananolab.nci.nih.gov
[20] www.elsevier.com/about/press-releases/research-and-journals/elsevier-and-us-national-cancer-institute-implement-reciprocal-linking-between-research-articles-and-datasets
[21] nanohub.org/groups/gng
[22] mgi.nist.gov/materials-data-repository
[23] mgi.nist.gov/materials-data-curation-system
[24] mgi.nist.gov/materials-resource-registry
[25] www.nano.gov/PublicWebinars
[26] www.nanoinformatics.org/

taken place since the founding of the NKI NSI. These workshops provide the primary venue where the nanoinformatics community gathers on a regular basis.

Strengthened by these interactions and Federal nanoinformatics activities, the community has expanded and matured in the past several years. Currently, an international group of scientists is developing an EU-initiated *Nanoinformatics 2030 Roadmap* that is slated for release in 2017.[27] This community-developed document is expected to describe the current state of the field, identify key gaps and needs, and outline possible short-, medium-, and long-term pilot projects that would address these needs. As representatives of agencies participating in the NKI NSI assess future directions, this roadmap may inform agency and signature initiatives activities in the context of other national priorities.

[27] The efforts to develop the *Nanoinformatics 2030 Roadmap* are being led by Dr. Andrea Haase (German Federal Institute for Risk Assessment). The U.S. contributions are being led by Dr. Fred Klaessig (Pennsylvania Bio Nano Systems, LLC).