

## **Nanoinformatics: A Means to Increased Collaboration in Research, Translation and Commercialization**

Dr. Martin Fritts

Contractor, National Institute of Standards and Technology  
and

Emeritus Researcher, Leidos Biomedical Research, Inc.  
Frederick National Laboratory for Cancer Research

University of Pennsylvania  
October 15, 2013

# Nanoinformatics 2013: Informatics for Nanomanufacturing

This talk addresses a means to achieve a goal – increased collaboration in nanotechnology research, translation and commercialization.

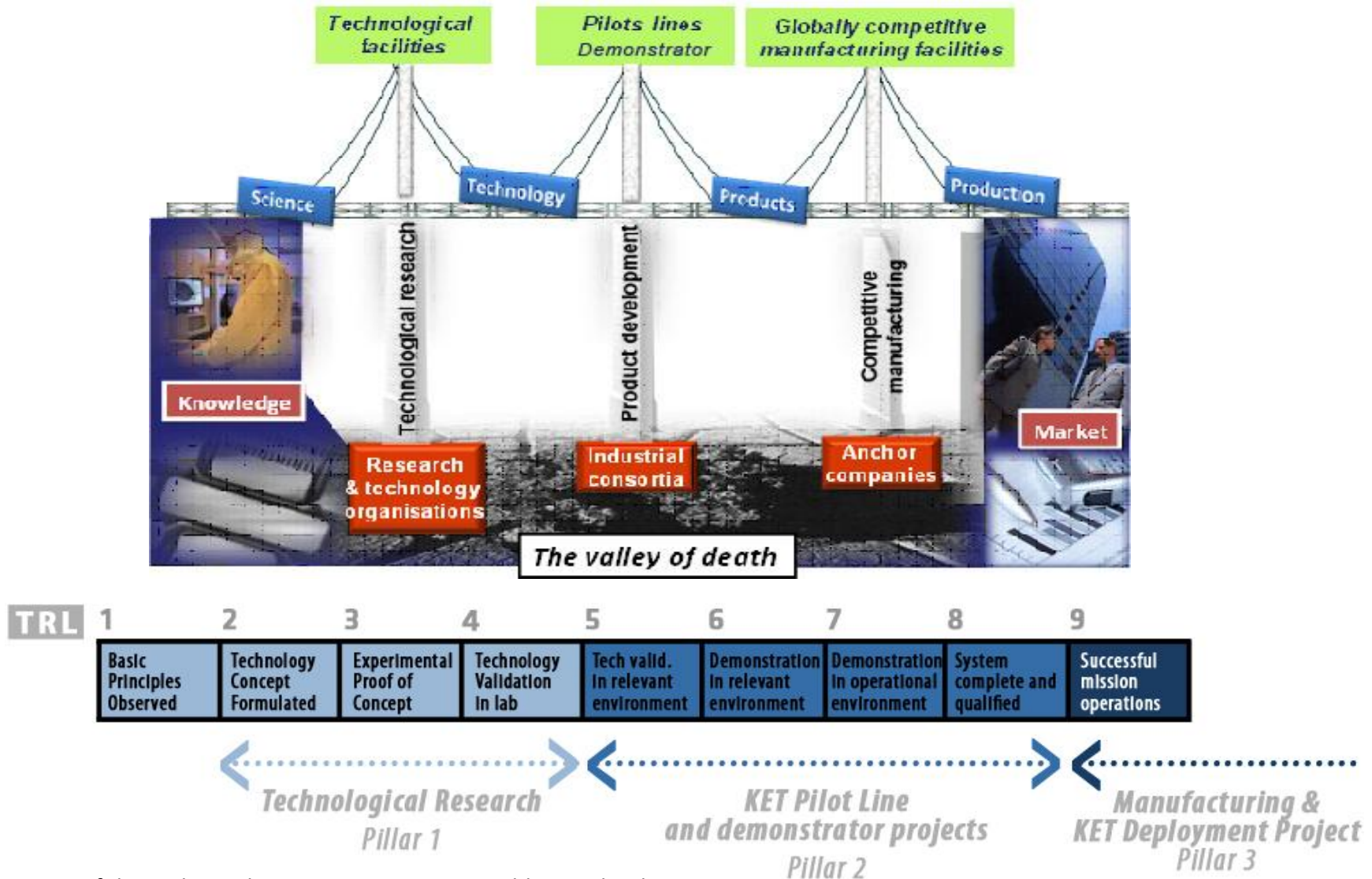
The NNI supports many different means to that same end - several of which will be mentioned as examples of current resources and practice. But informatics bridges those other efforts to amplify their impact.

This presentation is intended to be provocative:

- To cite opportunities to leverage other programs and initiatives,
- To suggest new ways to collaborate to accelerate commercialization while addressing EHS and ELSI issues, and
- In particular, to encourage increased inter-agency, inter-initiative, and industry collaboration in nanoinformatics – particularly in developing and initiating new pilots.

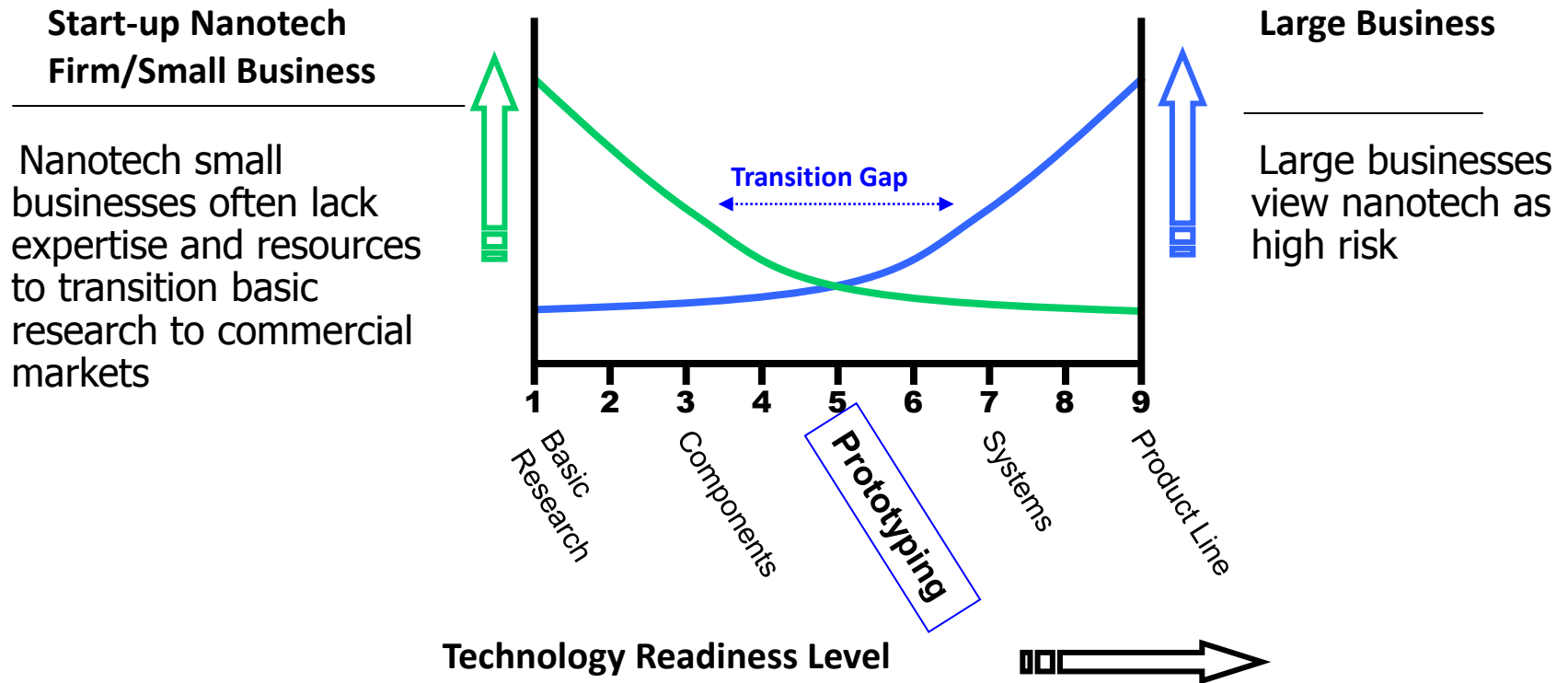
# Nanoinformatics 2013: Informatics for Nanomanufacturing

European integrated initiative to pass the "Valley of Death"



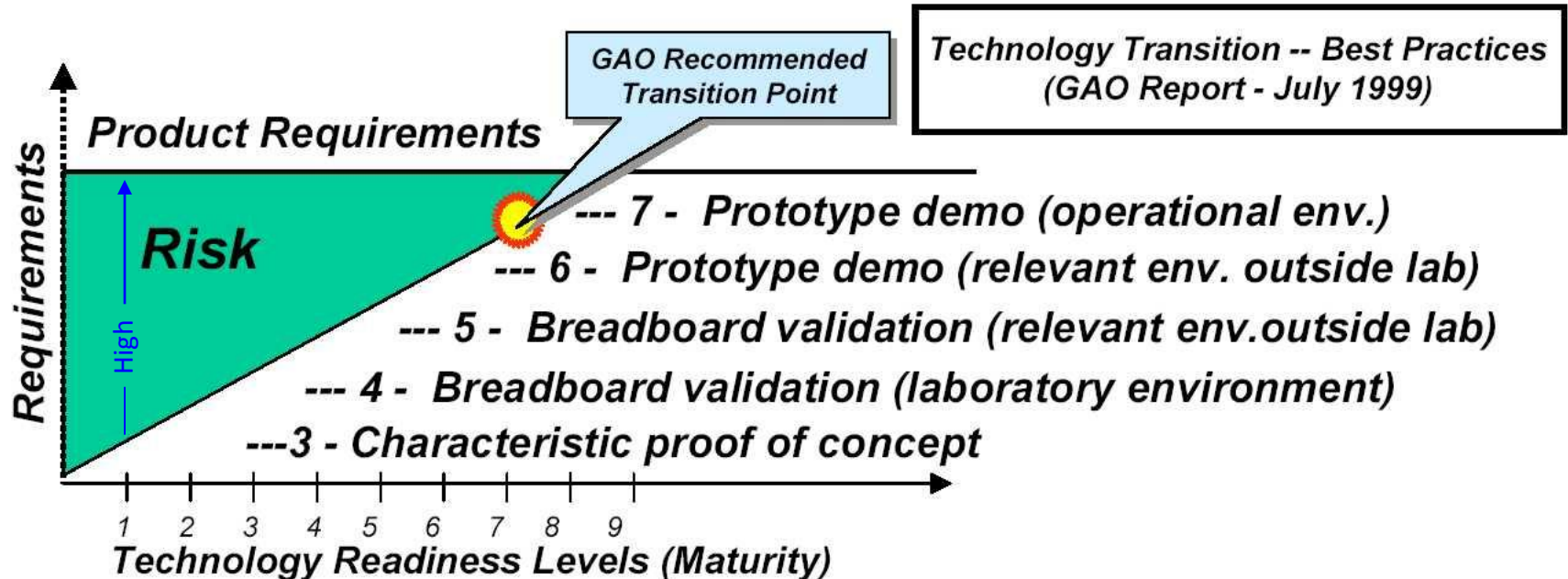
## Transition to Market

### Expertise & Resources



**There is presently a gap in transitioning nanotechnology from basic research to a commercial market, complicating the development of Nanotech products**

## Transition to Market - How and When

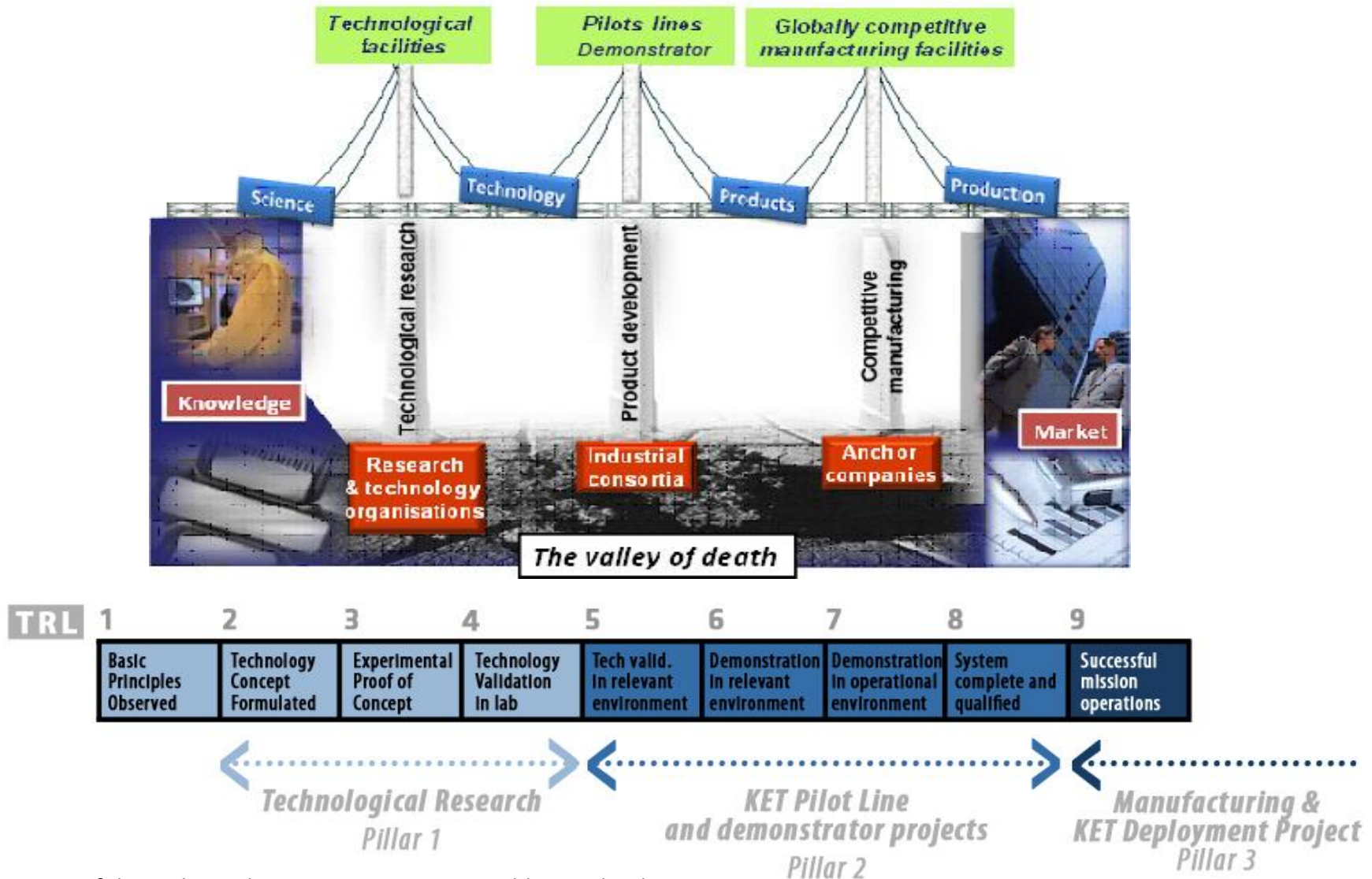


### Proof of Concept and Prototype demonstrations

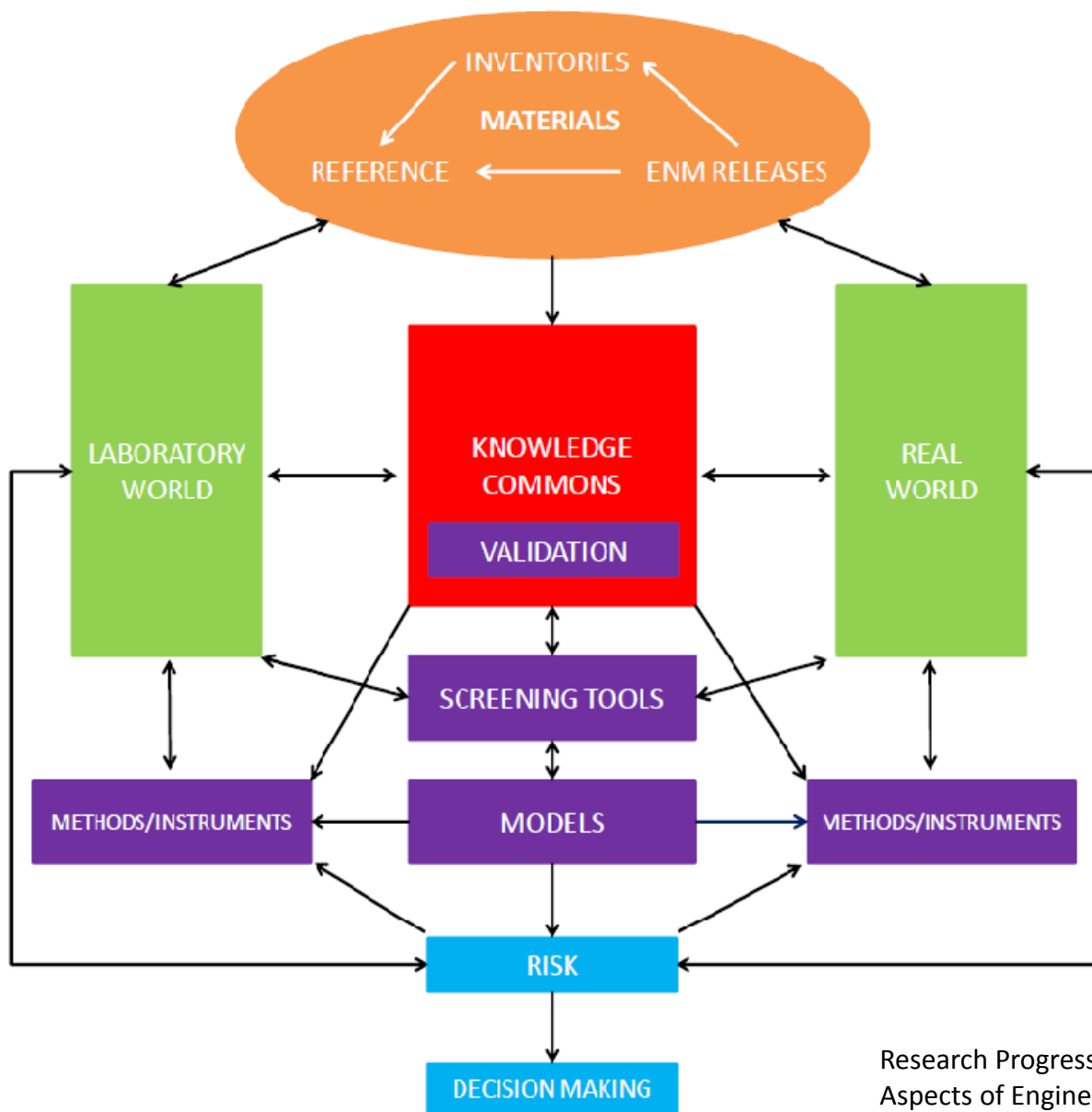
- Must ensure that the customer is a partner in product development
- Transition at TRLs less than 7 are high risk for sponsor/customer
- Bridge the “Valley of Death” for commercializing nanotechnology
- Provide credibility and accelerate transition to market

# Nanoinformatics 2013: Informatics for Nanomanufacturing

European integrated initiative to pass the "Valley of Death"



# Nanoinformatics 2013: Informatics for Nanomanufacturing



## The Nanotechnology EHS Research Enterprise

Production of ENMs  
Reference Materials  
ENM Releases

Locus of Collaborative  
Development of Methods,  
Models and Materials &  
Archiving, Curating, Sharing  
and Annotating Data

Lab Research to understand  
physical, chemical, and biological  
processes & mechanisms critical  
to assess hazard and exposure  
Real World Research to examine  
effects of ENMs on people and  
ecosystems

Methods, Models, Tools and  
Instruments needed for Research

Research Progress on Environmental, Health, and Safety  
Aspects of Engineered Nanomaterials, September 2013,  
NRC Report, National Academies Press

## *The Blind Men and the Elephant*



## A Comprehensive Knowledge Commons

Although the Knowledge Commons provides a framework for collaboration that could federate databases developed and managed by other entities, the list of relevant partners is long:

- The NNI consists of 28 federal partners and has developed over 100 R&D centers and user facilities.
- Partnering with centers and facilities developed by other initiatives such as the Materials Genome Initiative and the Big Data Initiative should also be considered.

The NSF and NIH now require management plans for data produced under their grants. That data, as well as information published in journal articles and supplementary information could also be included in a federated network.

## Pilot system Development and User Needs and Requirements

Informatics systems must fulfill user needs and requirements. As digital capabilities and the volume of data and information expand rapidly, requirements and needs evolve as well. Developing pilot systems and continually modifying those systems must involve all users, especially industry users, if the systems are to succeed.

The remainder of this talk discusses some of examples of pilot systems that need increased user input to be relevant, especially in nanotechnology, which seeks to design new materials and products for their lifecycle, including EHS and EdELSI considerations.

## 1. Assessing Data and Metadata Quality

Develop a framework for assessing data quality (model and experiment) based on data reproducibility and sufficiency of the metadata



### Data Readiness Levels (DRLs)

DRL 0. Invalid data

DRL 1. Raw or unscaled data

DRL 2. Scaled data

DRL 3. Data with defined precision or noise

DRL 4. Data with defined precision and noise

DRL 5. DRL 4 + data related to the larger body of scientific knowledge

DRL 6(X). Standards-quality data of X % measurement uncertainty

### Metadata Levels

**Poor:** Insufficient information: data cannot be reproduced/ interpreted by others

**Acceptable:** Others can reproduce and interpret the data; e.g., adequate descriptions of exp/comp methods used; descriptions of data formats.

**Excellent:** **Acceptable** + additional information; e.g., history/provenance, validation of the experimental methods and models.

## 2. The Nano-Silver Data Project

The Nano-Silver Data Project is a collaborative effort to structure and share the CEINT dataset with the Nanomaterial Registry and nanoHUB. The Nanomaterial Registry addresses such challenges as the need for standard methods, data formatting, and controlled vocabularies for data sharing. The Registry is an authoritative, web-based tool whose purpose is to simplify the community's level of effort in assessing nanomaterial data from environmental and biological interaction studies. All data-driven content is systematically archived and reviewed by subject-matter experts using a set of minimal information about nanomaterials (MIAN).

<https://www.nanomaterialregistry.org/>

[https://nanohub.org/groups/nano\\_silver](https://nanohub.org/groups/nano_silver)

<http://www.ceint.duke.edu/>

## 3. Data Curation/Sharing Workflows

NIST-Journal Cooperation to Improve the Quality of Published Experimental Data:

Pre-Acceptance Evaluation, On-line Tools,  
and IUPAC Recommendations

**Robert D. Chirico**,\*, Michael Frenkel, Joseph W. Magee, Vladimir V. Diky, Kenneth Kroenlein, Chris D. Muzny, Andrei F. Kazakov, Ilmutdin M. Abdulagatov, Gary R. Hardin, Theodoor W. de Loos, John P. O'Connell, Clare M. McCabe, Joan F. Brennecke, Paul M. Mathias, Anthony R. H. Goodwin, Jiangtao Wu, Kenneth N. Marsh, Ronald D. Weir, William E. Acree, Jr., Agilio Pádua, W. M. (Mickey) Haynes, Daniel G. Friend, Andreas Mandelis, Vicente Rives, Christoph Schick, Sergey Vyazovkin, and Ella Chen

Code: People at **NIST**, Journal Editors, Journal Managers  
Other important contributors: **Elsevier**, **American Chemical Society**, and **Springer** journal managers/staff

\* **Thermodynamics Research Center (TRC), (NIST), Boulder, Colorado**

## 3. Data Curation/Sharing Workflows

### **Facts leading to NIST-Journal cooperation...**

- Many published articles (~20 %) reporting experimental thermodynamic and transport property data contained significant numerical errors. (Reporting of nonsense uncertainties is not included in this number.)

The *rate of publication* of property data continues to increase rapidly. ( $\approx$  2-fold increase of data every 10 years.)

*Percentage of errors is increasing over time.* (Computers are great, but not always...)

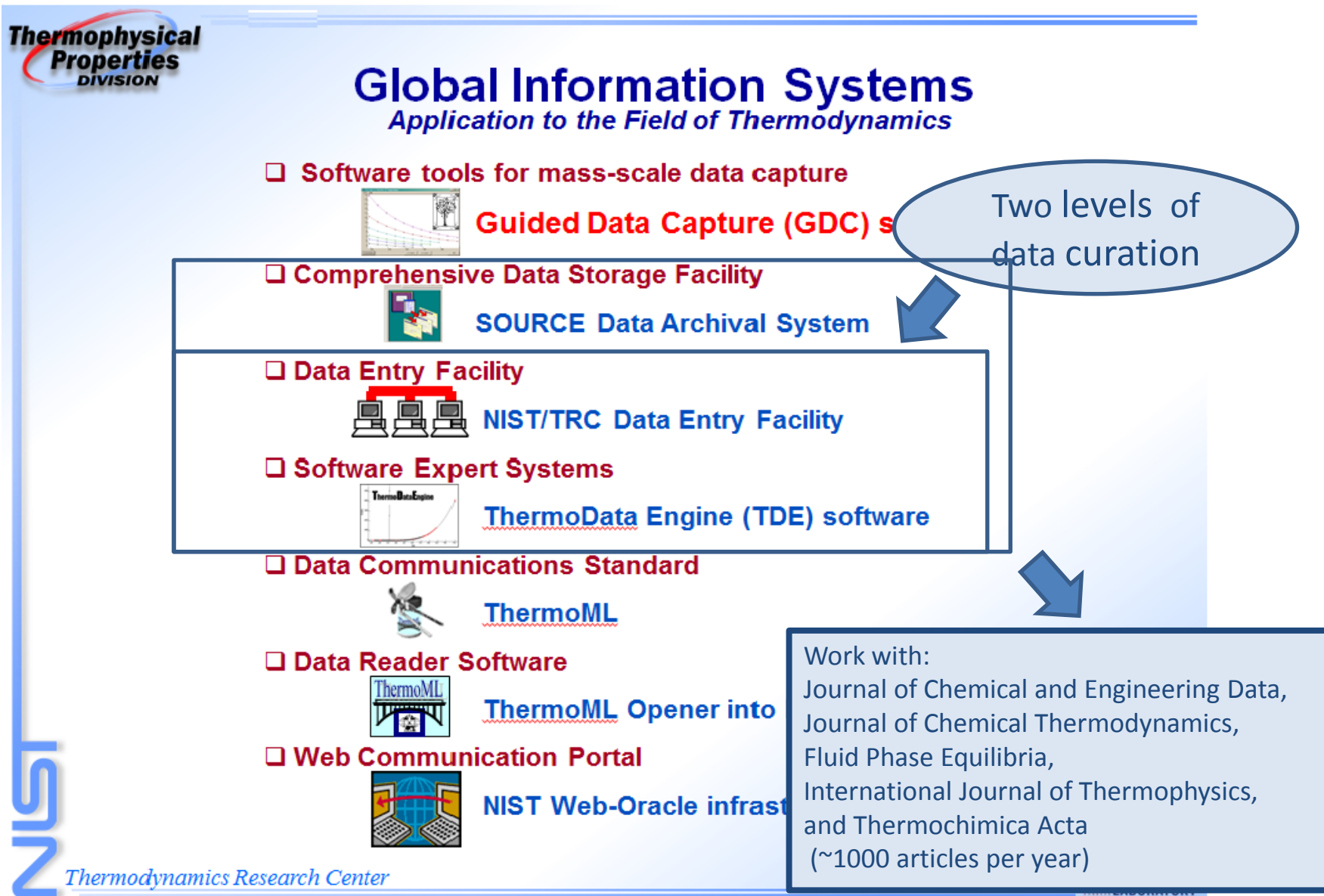
### **Result...**

- There are a lot of erroneous data in the literature... and the situation is getting worse.

### **Underlying problems...**

- **Problem 1:** Reviewers do not have the time or resources to *check* reported numerical data *against available literature data*.
  - **Problem 2:** Reviewers do not have the time or resources to check the *quality of literature searches* by authors.
  - **Problem 3:** Tabulated data are *very rarely plotted* at any time in the review process.
    - This would reveal many problems.
- \* The implemented procedures are designed to help with all of these problems.**

## 3. Data Curation/Sharing Workflows



## 3. Data Curation/Sharing Workflows

**Background to what we do within the  
NIST Thermodynamics Research Center**

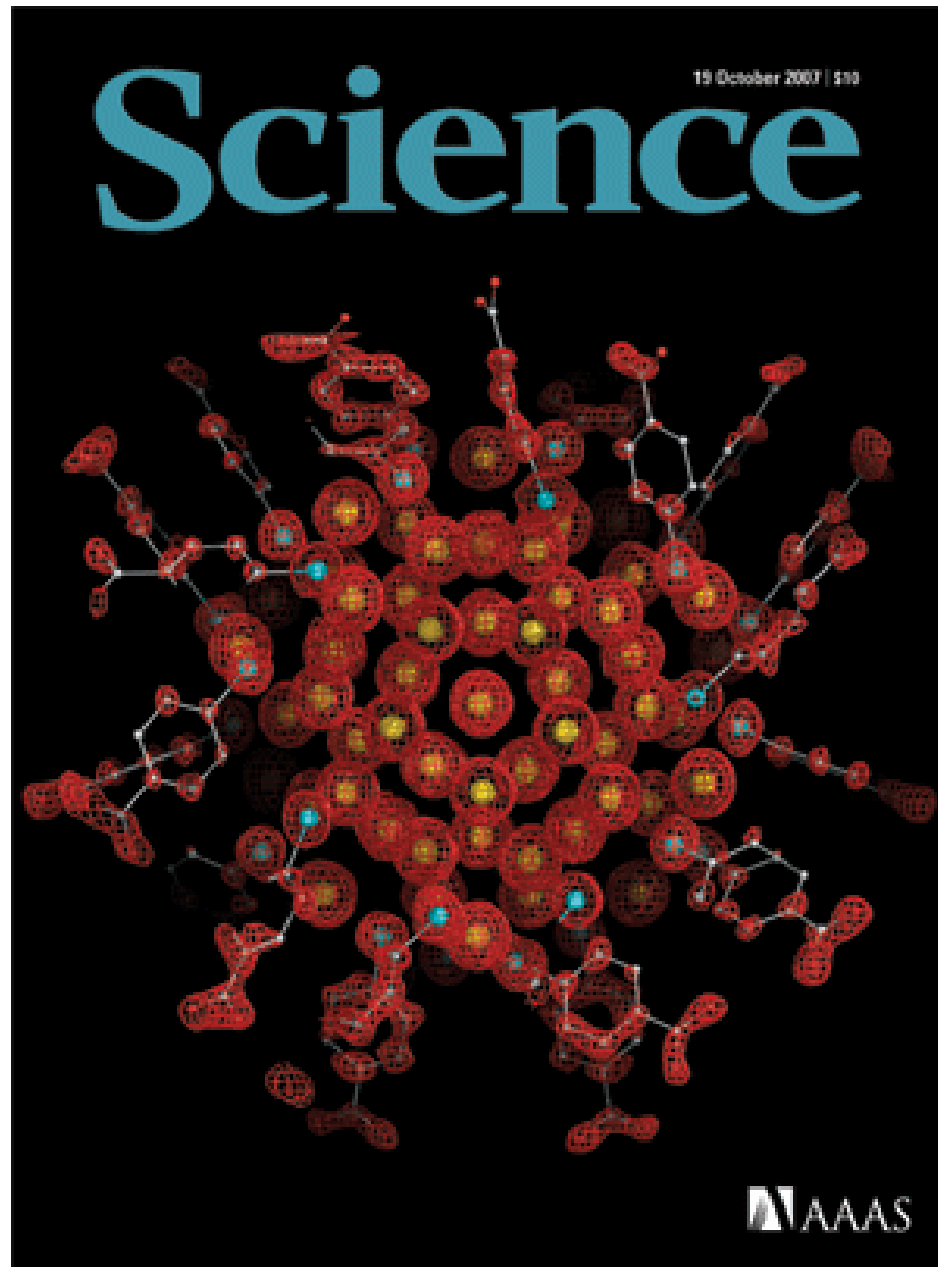
**Goal/Mission:** Provide *critically evaluated* thermophysical and thermochemical property values of chemicals (and mixtures) for use by industry, academia, and other government agencies for...

- Chemical process development & optimization
- Fundamental research into molecular properties
- Regulatory decisions
- Many others

## 4. Better Standard Methods More Quickly

ASTM's E56 Committee on Nanotechnology is considering initiating new processes for more efficient development of standard test methods:

- Video protocols will be used to provide more detailed information on protocol details which are now inadequately described, such as sample preparation;
- High throughput screening techniques should be used to quantify sensitivities of the protocol parameters;
- Informal tests of the video protocol using provided study materials will allow review and comment by interested laboratories using web-enabled applications;
- Once a consensus among the reached on the video, the standard document will be developed and voted upon, hopefully attaining consensus more quickly;
- A formal interlaboratory study will be conducted to provide the necessary measures of the error and uncertainty to be expected with the protocol as required in ASTM standard test methods;
- The video will be available with the standard for training and certification.



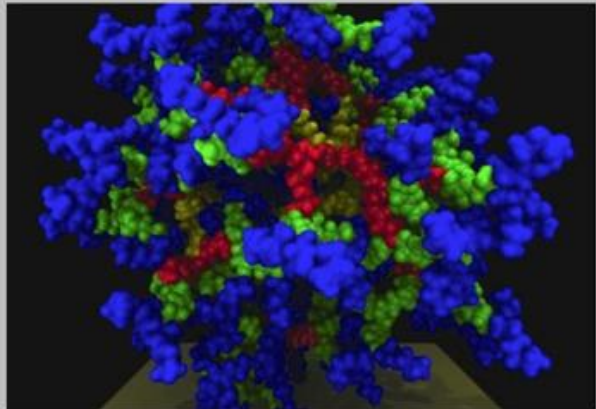
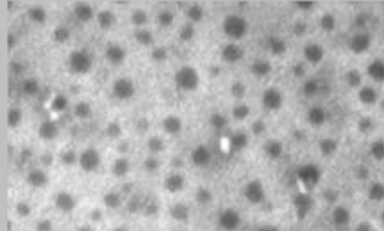

## 5. Validating and Sharing Material Models – and Predictive Models

# Nanoinformatics 2013: Informatics for Nanomanufacturing

## Collaboratory for structural Nanobiology

Advanced Structure Analysis Collaborators Services

Welcome 2013







*"Our lack of adequate in vivo predictive capabilities are the major source of failure in the progression from the research and development phase to clinical use."*

**-NCI caNanoPlan**

Enter

Raul.Cachau@NIH.GOV



Privacy Policy About us Services Science

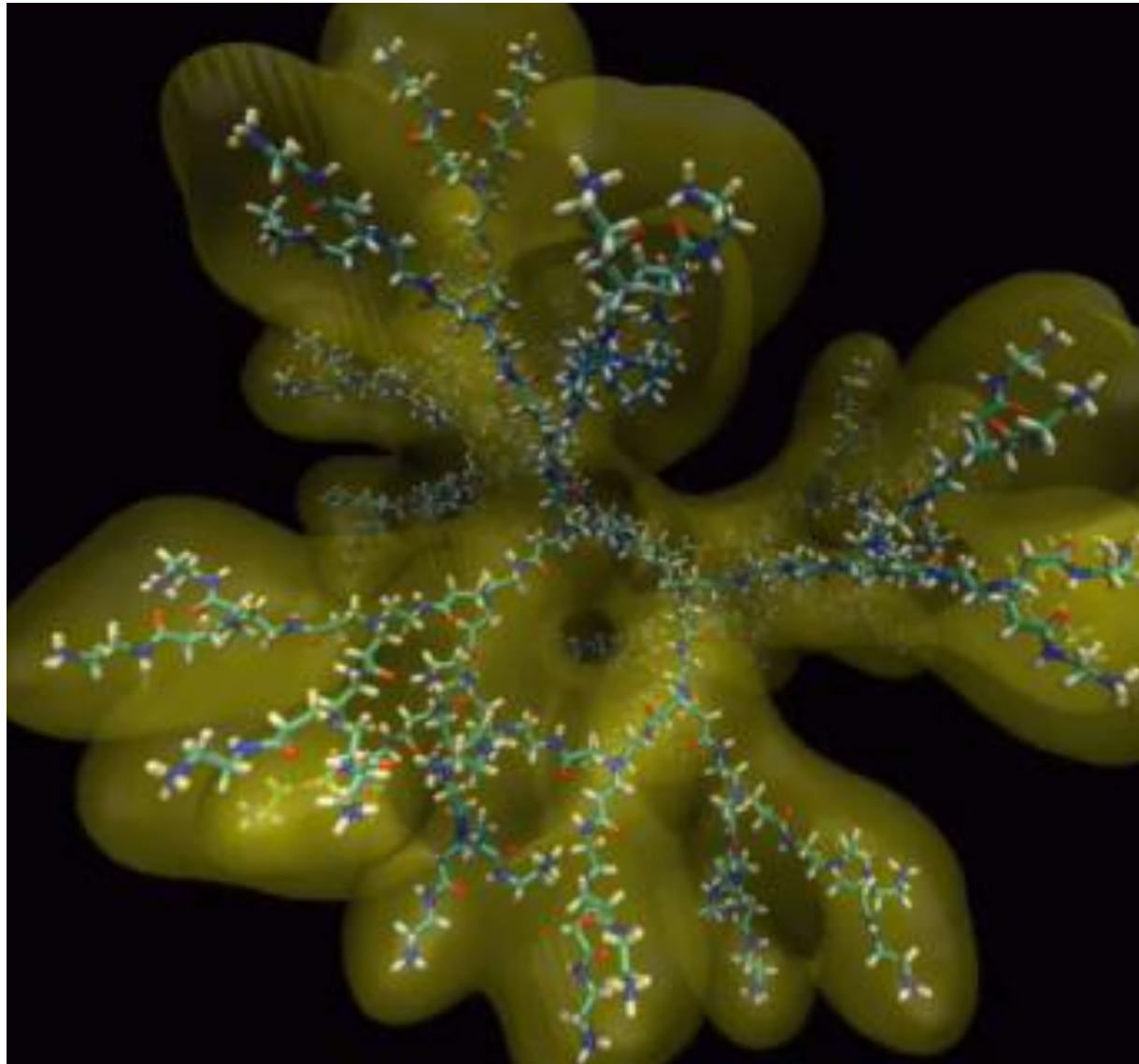
This work has been funded in part with funds from the NCI-NIH (Contract HHSN261200800001E). The contents of this publication do not necessarily reflect the views or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Play

# Nanoinformatics 2013: Informatics for Nanomanufacturing

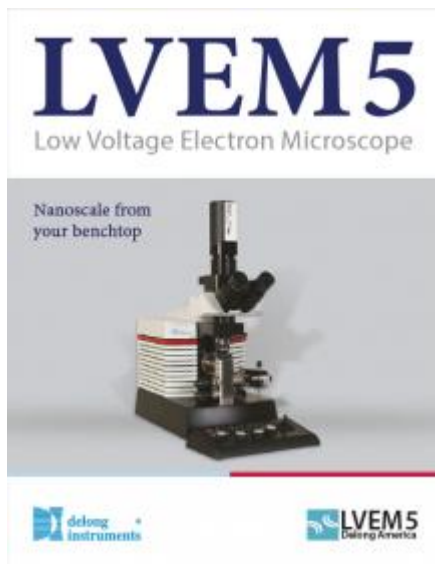
## Collaboratory for structural Nanobiology





## 6. Instrumentation Development

We are in the midst of a revolution in instrumentation capability and cost – not just for research facilities, but also for the factory floor. The placement of these instruments at nanotechnology user facilities could aid in both familiarizing future customers with their capabilities, but also eliciting user requirements .



### **Low-voltage Electron microscopes**

The DeLong low-voltage EM instrument has a two-foot height, 2nm resolution, high contrast for soft material, SEM, TEM and STEM modes, and can be transported easily , and used in a work floor environment and in the field.

## 6. Instrumentation Development

We are in the midst of a revolution in instrumentation capability and cost – not just for research facilities, but also for the factory floor. The placement of these instruments at nanotechnology user facilities could aid in both familiarizing future customers with their capabilities, but also eliciting user requirements .

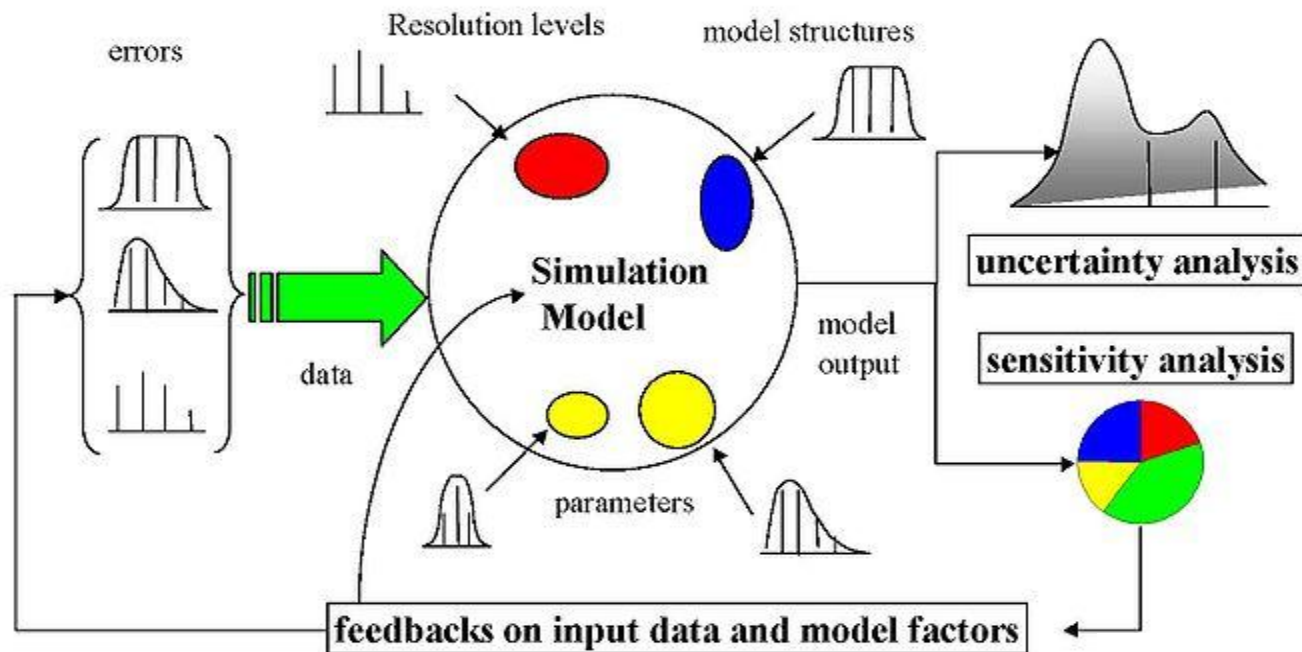
### **Bright spatially coherent synchrotron X-rays from a table-top source**

State-of-the-art X-ray sources can now produce coherent high-brightness X-rays of greater than kiloelectronvolt energy and promise a new revolution in imaging complex systems on nanometre and femtosecond scales. ...Here we demonstrate the use of a new generation of laser-driven plasma accelerators, which accelerate high-charge electron beams to high energy in short distances, to produce directional, spatially coherent, intrinsically ultrafast beams of hard X-rays. **This reduces the size of the synchrotron source from the tens of metres to the centimetre scale**, simultaneously accelerating and wiggling the electron beam. The resulting X-ray source is 1,000 times brighter than previously reported plasma wigglers and thus has the potential to facilitate a myriad of uses across the whole spectrum of light-source applications.

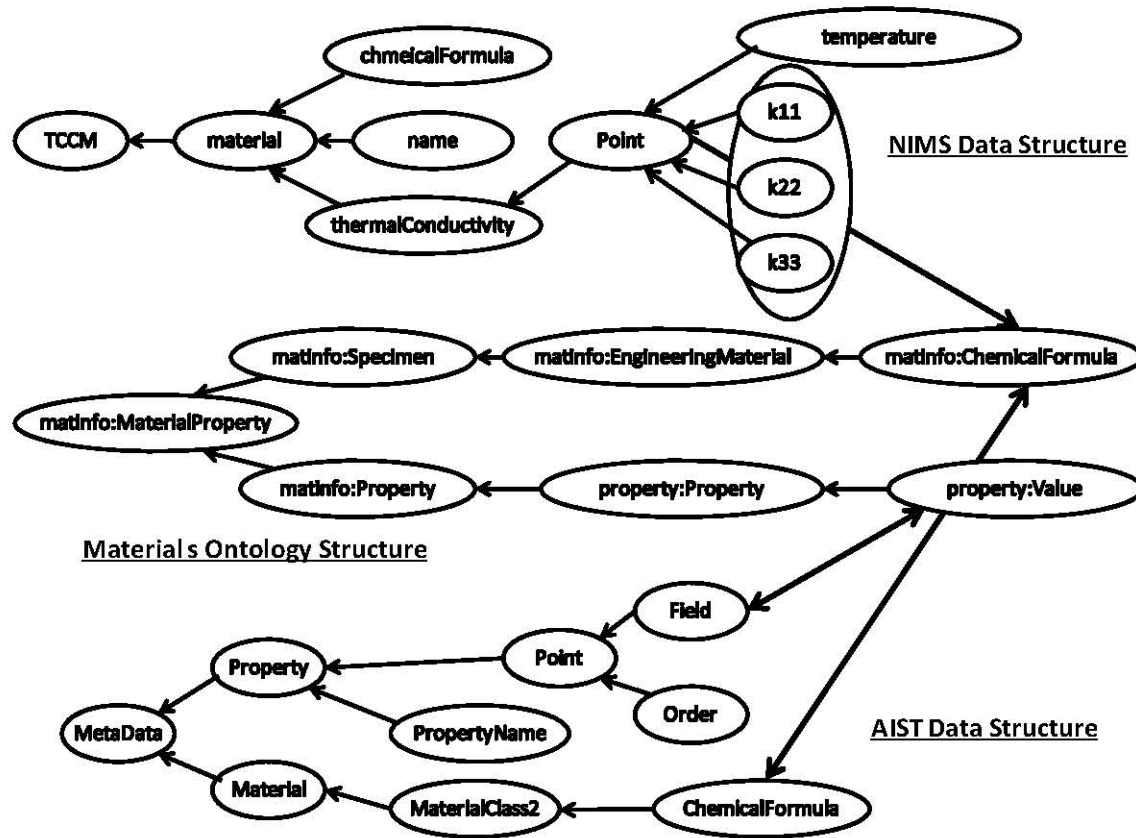
S. Kneip, C. McGuffey, et.al., Nature Physics 6, 980 (2010). doi:10.1038/nphys1789

## 7. Uncertainty and Sensitivity Analysis

Data analysis and experimental design –  
especially with high throughput screening  
Risk analysis and deliberative risk management  
Nanomaterial design and process design



## Mapping Ontologies & Metadata



Data schema structures for thermal conductivity of the NIMS, Materials Ontology, and AIST databases and correspondence of their data fields; “matinfo:” is the prefix of the material information ontology.

## Collaborate with Other Programs

### *Enable & Enhance Exchange*

- Develop and deploy repositories
- **Develop and disseminate materials informatics infrastructure**
  - Enable data discovery through tools and standards
  - Capture data from scientific workflows and archival sources
  - Engage with stakeholders to determine needs and disseminate best practices
- Integrate across length and time scale
- Build and Test infrastructure through Pilots

MGI James Warren 28/6/2013

Tuesday, May 28, 13

[http://www.nitrd.gov/nitrdgroups/index.php?title=Data Sharing and Metadata Curation: Obstacles and Strategies](http://www.nitrd.gov/nitrdgroups/index.php?title=Data_Sharing_and_Metadata_Curation:_Obstacles_and_Strategies)

(Big Data Program , with agencies in Networking and IT R&D Program- NITRD)

## On the Shoulders of Giants...



Cedalion standing  
on the shoulders of  
Orion

from  
*Blind Orion  
Searching for the  
Rising Sun*

by  
Nicholas Poussin,  
1658

*"Bernard of Chartres used to say that we are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size."*

John of Salisbury, 1159

## Summary

Sir Isaac Newton alluded to the essential collaborative nature of science in 1676, noting that he stood “on the shoulders of giants”. This talk presented some recent efforts to employ informatics tools and applications to improve collaboration on nanotechnology research, development and translation among multidisciplinary, multi-agency partners.

- 1) increasing the availability, **reliability and reproducibility of data** on the properties of nanoparticles by accelerating the development of standard protocols, interlaboratory studies, data curation methodologies and access to improved instrumentation;
- 2) providing **semantic search** for data discovery and sharing based on common vocabularies, ontologies and mapping tools; and
- 3) **collaboratories** for developing, validating and sharing molecular models of nanomaterial, labeling structural motifs, and collaboratively developing predictive models for nanomaterials, their interaction with biological environments, the evaluation of their EHS risk, and incorporation of input concerning the educational, ethical, legal, and societal implications of the research, development and commercialization efforts.

Please help!

Thank You

## Attempt new ways to use existing mechanisms for more effective collaboration

Example: Many initiatives are supported by multiple agencies. However, agencies must use their funding within the bounds of their mission statement, and specifically not to support another agency's mission. As a result, infrastructure to support common needs (agency or industry) is more difficult than it needs to be.

However, FFRDCs were instituted to let agencies benefit from business efficiencies and can consider “support for others”. Driving infrastructure development through FFRDCs could resolve the COI problems for interagency support.

## Consider Using Social Media with Database Tools

- “I’ll give you my data if I can send it through Facebook”
- LinkedIn + dating service + selective anonymity + strict governance could provide a means to gradually establish a relationship for:
  - finding expertise, by discussing need and capability before revealing company identity
  - Begin discussions of IP portfolios anonymously, gradually leading to signing of NDAs and open negotiation
  - Find suitable partners for commercialization, ...

## The “Measurement Equation”

- The true but unknown value of the measured quantity or **response**, denoted by **R**, is related to the true but unknown values of the **arguments**

$$(\alpha_1, \dots, \alpha_k)$$

Key parameters

by a **known relationship (i.e., function)**.

$$R = f(\alpha_1, \dots, \alpha_k).$$

System Sensitivities

- Expanding to **first order** via a Taylor Series expansion

$$R(\alpha_1, \dots, \alpha_k) = R(\alpha^0) + \sum_{i=1}^k \left( \frac{\partial R}{\partial \alpha_i} \right)_{\alpha^0} \delta \alpha_i = R^0 + \sum_{i=1}^k S_i \delta \alpha_i,$$

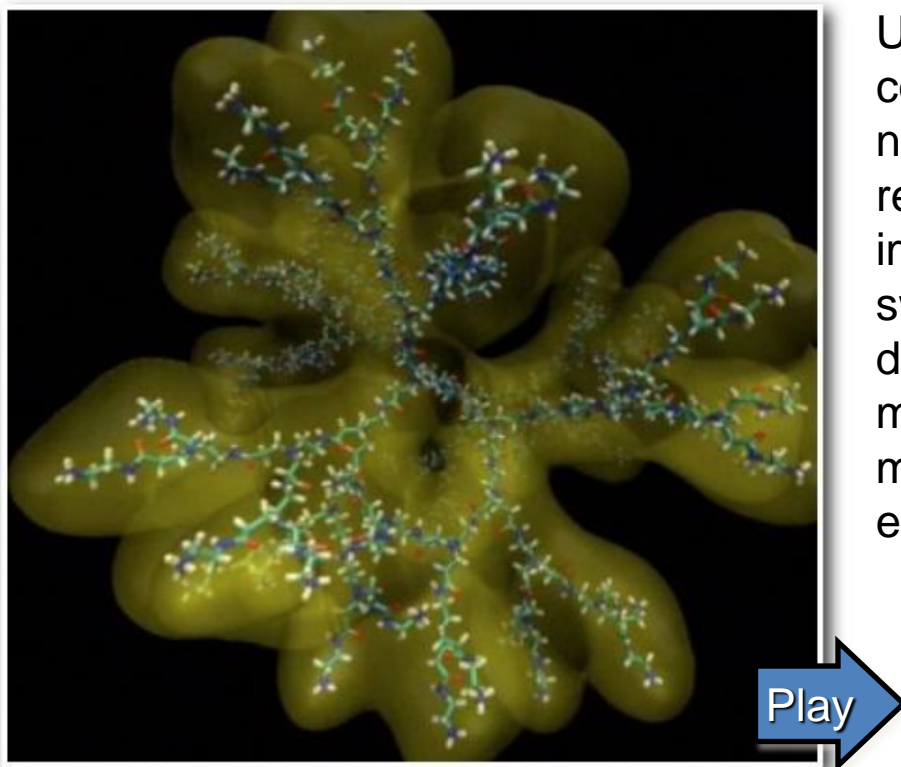
Errors

- This “**measurement equation**” can be interpreted to represent not only results of indirect measurements but also *results of computations*.

# Collaboratory for structural Nanobiology



## Nanotoxicity - SAR - Drug Design



Understanding the mechanisms that control the properties of engineered nanomaterials (ENPs) at the atomic resolution is critical for the design of improved ENPs. The insert shows the swelling process of a PAMAM dendrimer. The entire process was modeled using quantum mechanical methods for the description of the entire particle.