Nanoinformatics and the nanomaterial research community:

Strategies for Sharing and Convergence

Thursday 27 March, 2018
NanoWG

Christine Ogilvie Hendren
Jaleesia Amos

Camille de Garidel-Thoron

Claus Svendsen
Tassos Papadiamantis

Iseult Lynch,
Marianne Matzke

The Commons

NanoInformatics Knowledge Commons

CEINT

eNM eNanoMapper

Serenade

NIKC

NanoFASE

iCEINT

us-eu
bridging nanoEHS research efforts

What is needed from nanoinformatics?

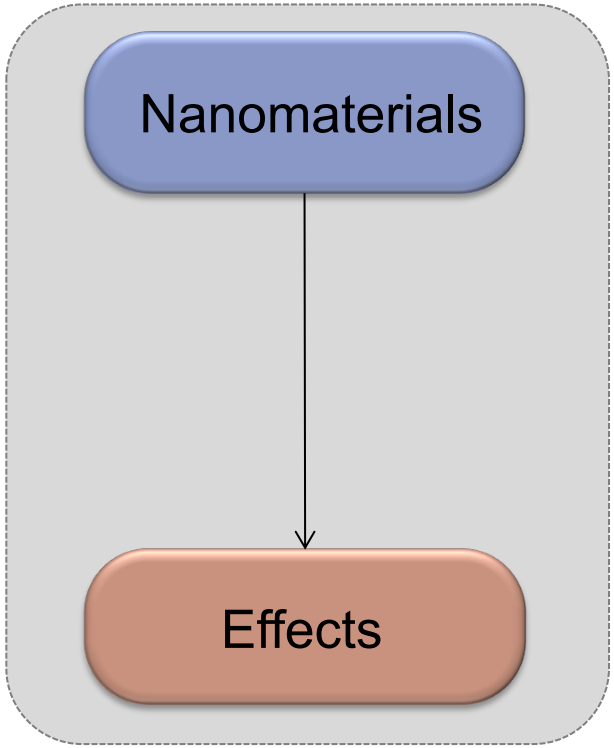
1. To serve the combined data and analytical needs of regulators, researchers, and industry, we need to:
   1. Combine and mine emerging datasets within and across projects to enhance scientific discovery
   2. Harvest knowledge from previous studies for comparative analyses
2. To functionally make this progress happen across multiple parties, we need to:
   1. Get the scale, sequencing, communication, and incentives right
   2. Create varying sharing practices to allow balance between openness and privacy

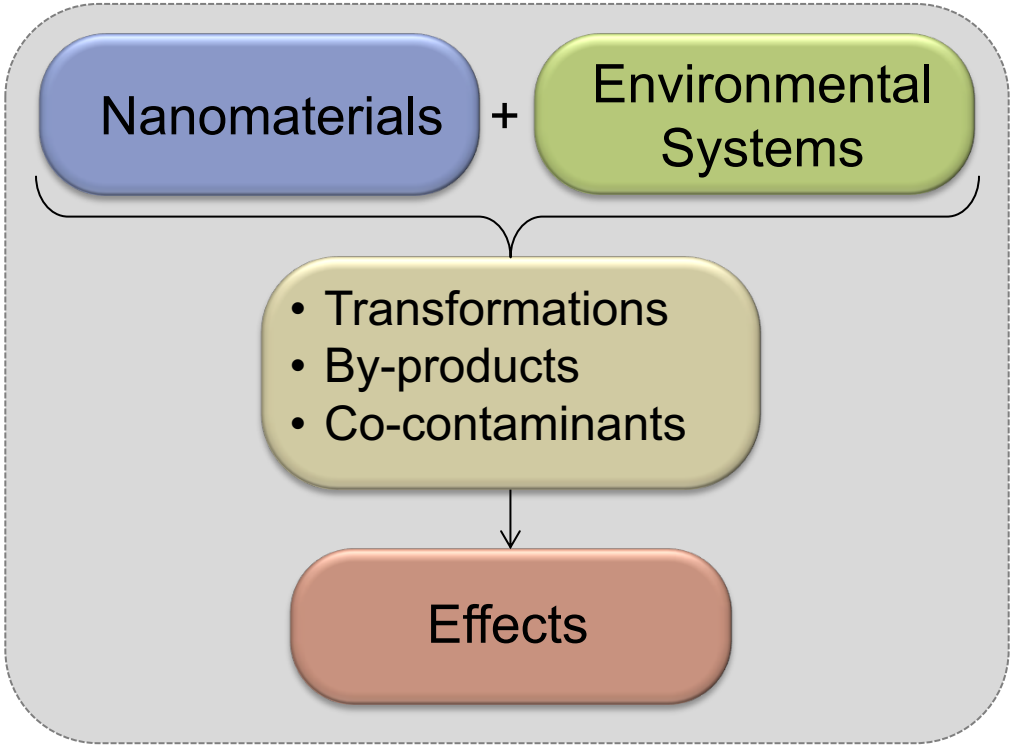
What can research projects deliver?

Building infrastructure and standards from 3-4 year projects with focused research missions requires:

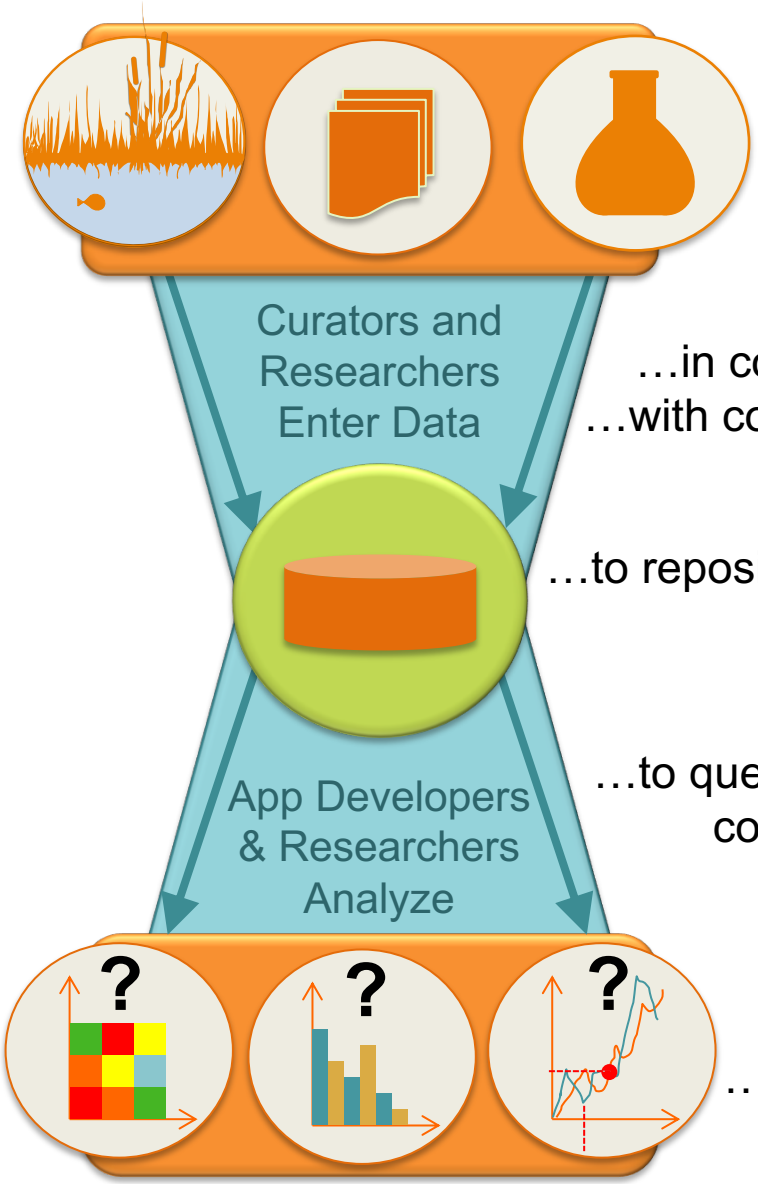
- Realistic expectations
- Modular approach to allow stepwise progress
- An open and shared approach towards a common goal

Environmental Complexity: A Necessary Reality





How to Address the Complexity?



Collect data

…in consistent format
…with consistent language

…to repositories

…to query, combine, compare

…to answer research questions

…to pose new research questions

…to make sure we get enough of the same data that we can compare

Start Focused and Small, Generalize Later

e.g. NanoInformatics Knowledge Commons (NIKC) Approach:

What are the key research questions? What is needed to answer them?

Can **key functional assays** in a reference system predict (a) exposure, (b) biouptake and (c) toxicity?

- What reference systems?
- What parameters to characterize?
- What comparisons to make?

Leveraged Projects and Programs

NBI KNOWLEDGEBASE

NANOMATERIALREGISTRY

ENM eNanoMapper

ISA-TAB-Nano

us-eu bridging nanoEHS research efforts

- Established ontologies
- Existing file-sharing templates
- Instance of characterization
- Test datasets

NIKC

Many Projects, One Repository: A Pilot Integration

NIKC

- External datasets
- Expanded dictionaries
- Feedback & design requirements

- Combined analytical power
- Templates for more complete data collection

Active collaborations: US & EU

NanoCommons
Nano-Knowledge Community

CEINT

NanoFASE

Serenade

rivm
National Institute for Public Health and the Environment

Mod Nano Tox

NanoMILE

NANSREG

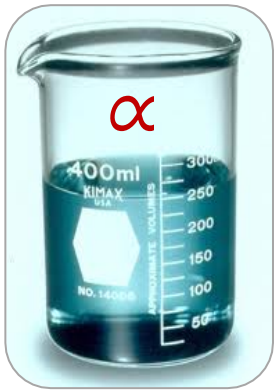
ProSafe

- Current Parterships:
  - Focused on environmental and fate issues
  - Rich data and research interest overlap
  - Desire to develop common standards at outset
    - Protocols
    - Ontologies
    - Modular, flexible updating based on targeted community stakeholder input

- Each relationship requires:
  - Custom template development
  - Iterative communication technical leads

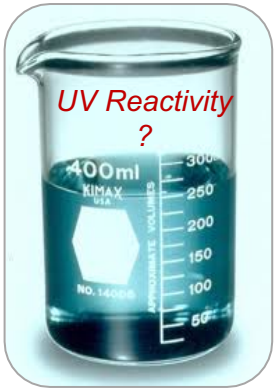
Focus on Shared Methods & Questions

Functional Assay Methods and Data Input Tools for Key Parameters

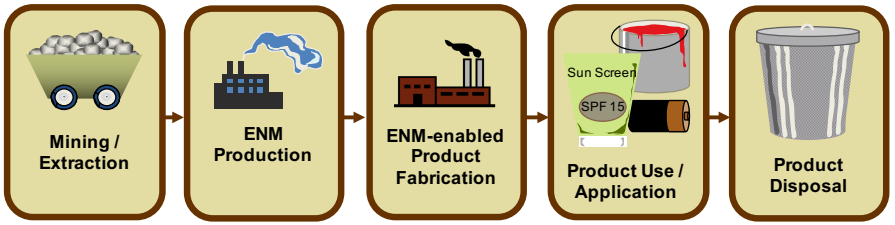


Established

Developing



Next?



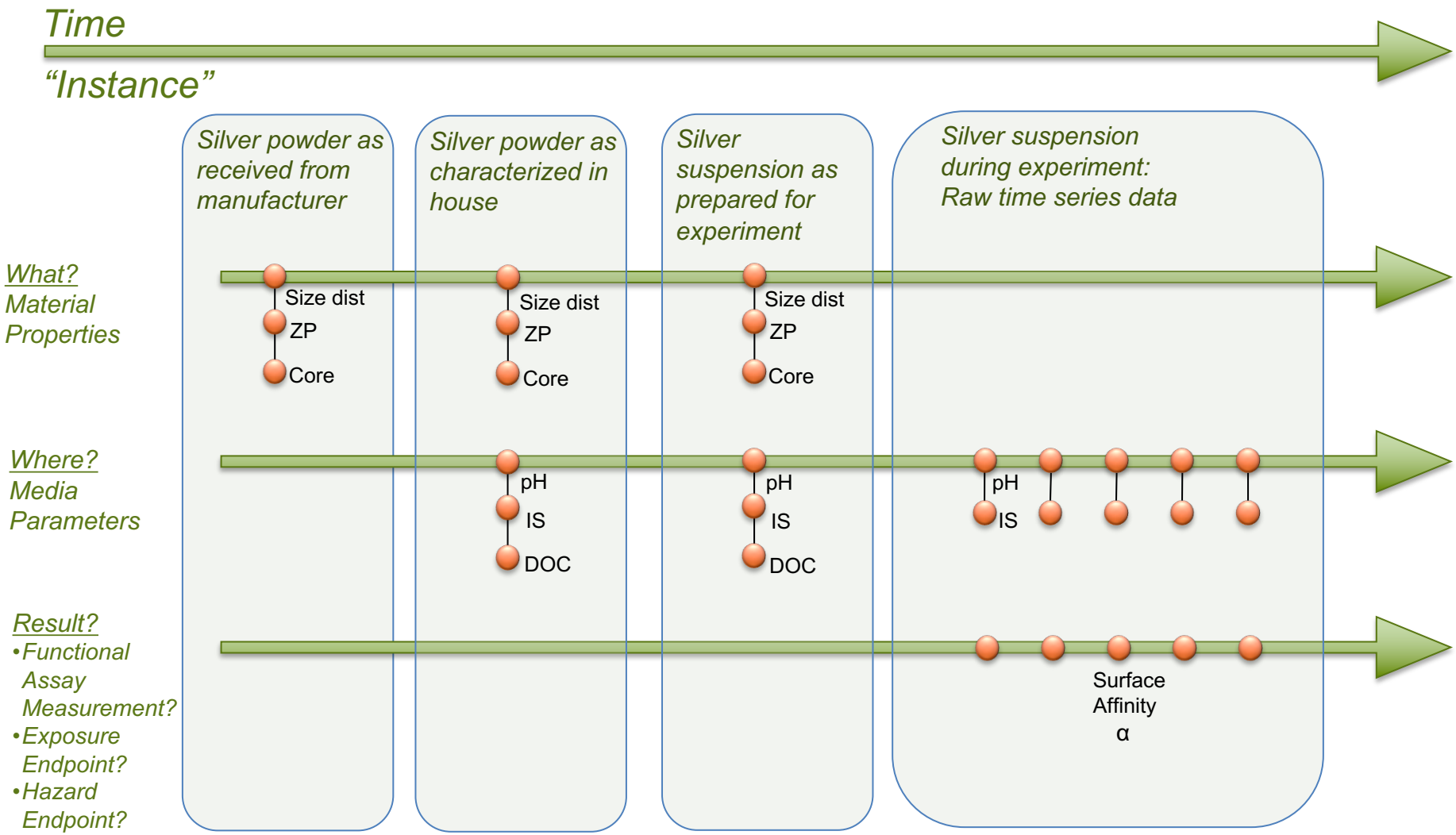
Nanomaterials + Environmental Systems

Exposure Descriptors
- Surface affinity ∝
- Dissolution
- Other transformations

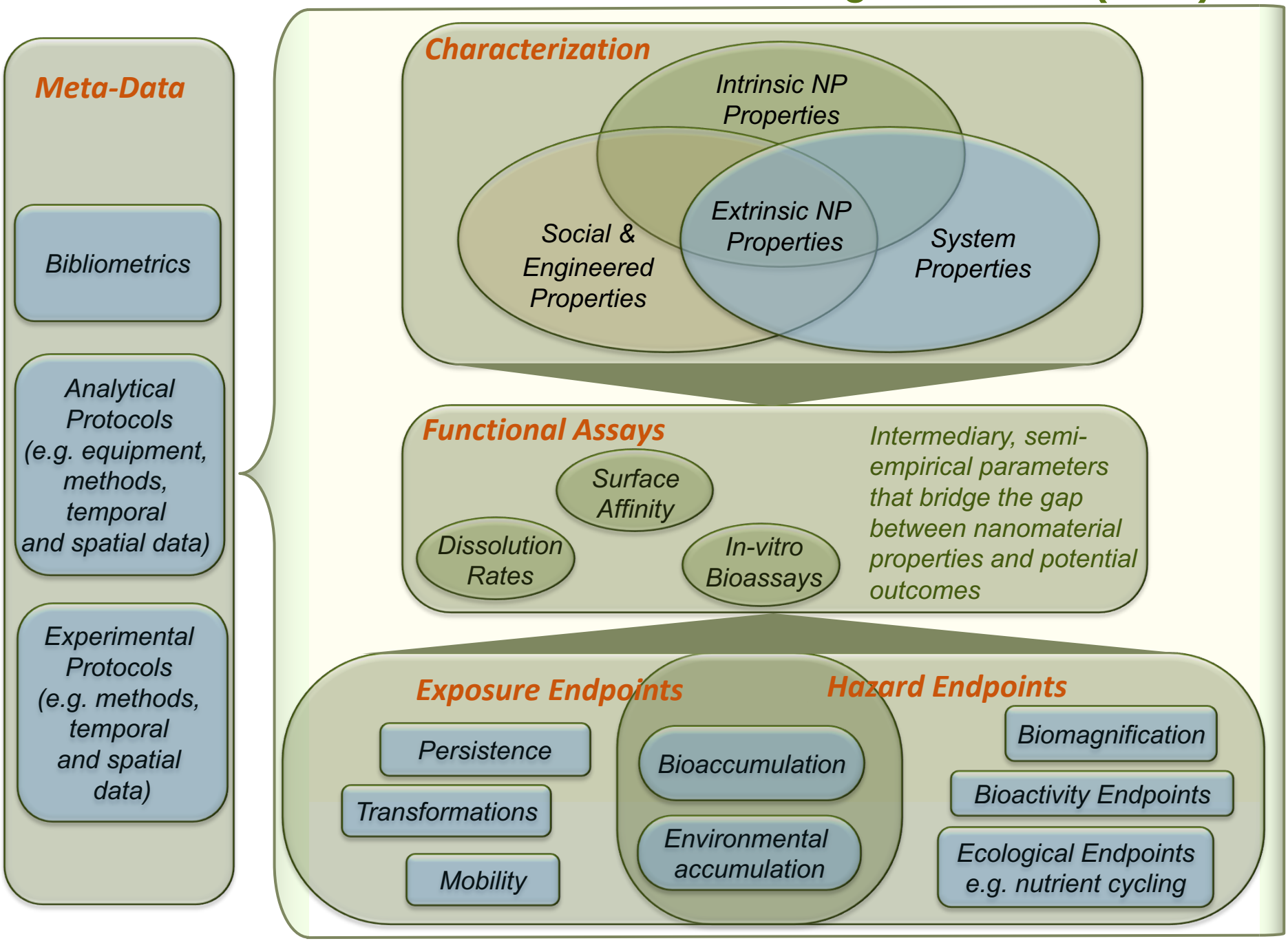
Effects

- Parameter modules added one by one via:
  - Utilizing existing ontologies & augmenting NIKC dictionaries with new terms only when needed
  - Custom template development

Why We Need to Capture Temporal Data to Understand Nanomaterial Fate, Exposure, Hazard and Risk



Data in the CEINT NanoInformatics Knowledge Commons (NIKC)

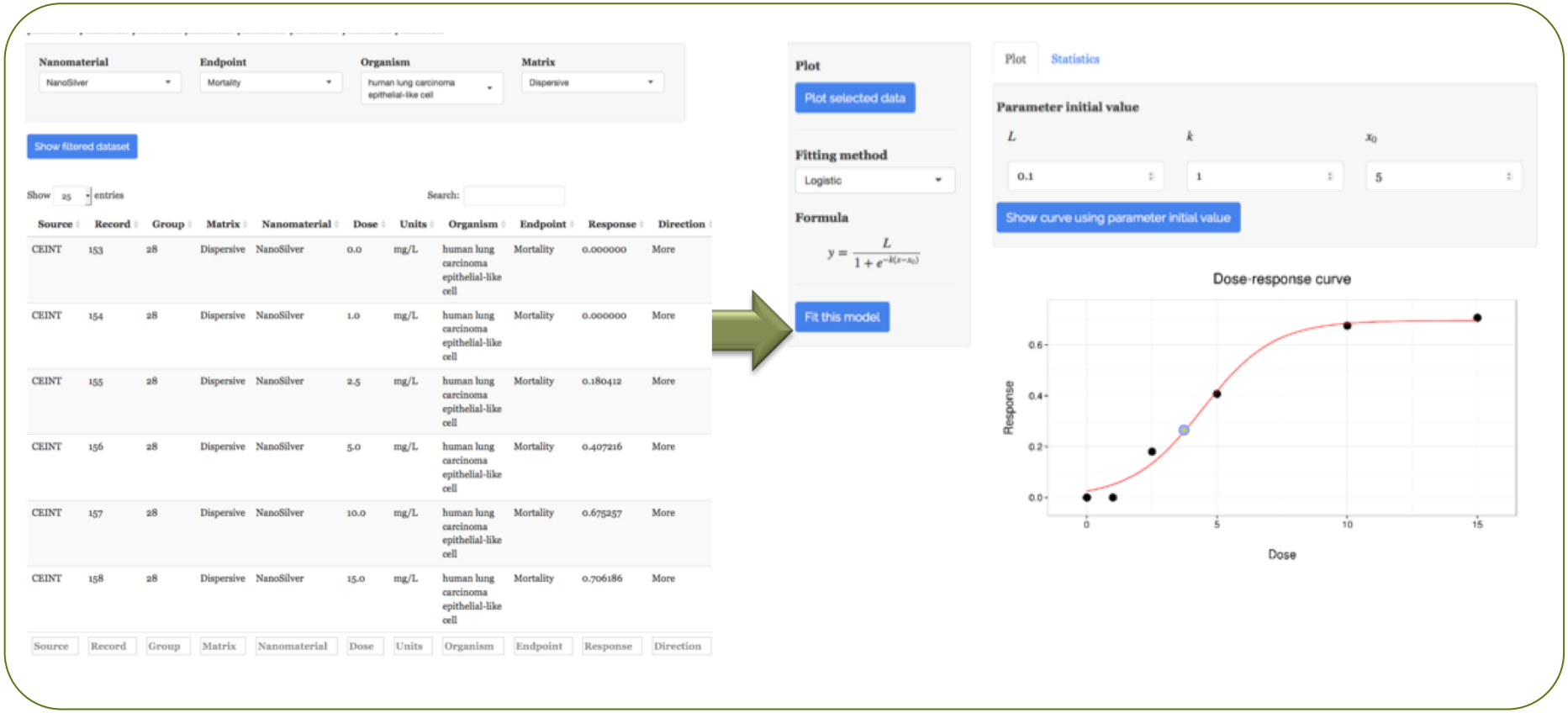


Why Would Researchers Self-Curate Their Raw Experimental Data?:

Improved Insight from Viewing Data in the Context of the Commons

e.g. NanoPHEAT (Nano-Product Hazard & Exposure Assessment Tool)

- Interactive Tool Combining Literature, experimental data, and model estimates
- Superimposes estimated exposures from nano-enabled products onto pristine material dose-response curves



Building a Research-Based NanoInformatics Knowledge Commons That Serves Risk Assessment Information Needs

Have	Developing	Need
Data exchange formats, piloted and adapted across EU and US (ISA-TAB-nano)	Pilot collaborative dictionaries linked to existing ontologies	Integrated, federated dictionaries linked to ontologies to provide "drop-down" values
Established curation templates and training documents	Early tools for researcher-driven curation of functional assay data	Continued development of custom web-enabled tools for researcher-driven curation
Established cyberinfrastructure (NIKC)	Shared US-EU research community infrastructure (through NanoCommons)	Long-term sustained and actively maintained public infrastructure (e.g. EUON)
Pilot query and visualization tool	Cross-project query and visualization tool	Stakeholder-driven development of web-enabled tools for analysis and visualization
Pilot collaborative core of teams combining data across projects into NIKC	Communal tools for real-time updates to terms, protocols, and data sets	Community charter with standard methods to facilitate user-driven updates of ontologies, tools and structures

MANUAL → *AUTO*