

Annie M. Jarabek

US EPA National Center for Environmental Assessment

ARA Beyond Science and Decisions Workshop

November 2, 2012

**ADVANCING MULTI-SCALE INTEGRATION OF
HUMAN HEALTH AND ENVIRONMENTAL DATA:
COMPUTATIONAL AND CONCEPTUAL
INTEROPERABILITY**

Disclaimer

- These views are those of the author and do not represent US EPA policy.

Overview

- ⦿ Background
 - Interoperability and multi-scale modeling
 - US EPA ORD Global to Genome (G2G) project
 - SOT CCT B4BD Workshop
- ⦿ Recommendations
- ⦿ Draft system design description
- ⦿ Next steps



Interoperability

- ⦿ The ability of two or more systems or components to exchange information and use the information in end user applications, present or future, any restricted access or implementation
- ⦿ The ability of diverse systems and organizations to work together

LCIM: Levels of Conceptual Interoperability Model

Wang et al. (2009)

Composability

6: Conceptual

Modeling /
Abstraction

5: Dynamic

4: Pragmatic

Simulation /
Implementation

3: Semantic

2: Syntactic

Integratibility

1: Technical

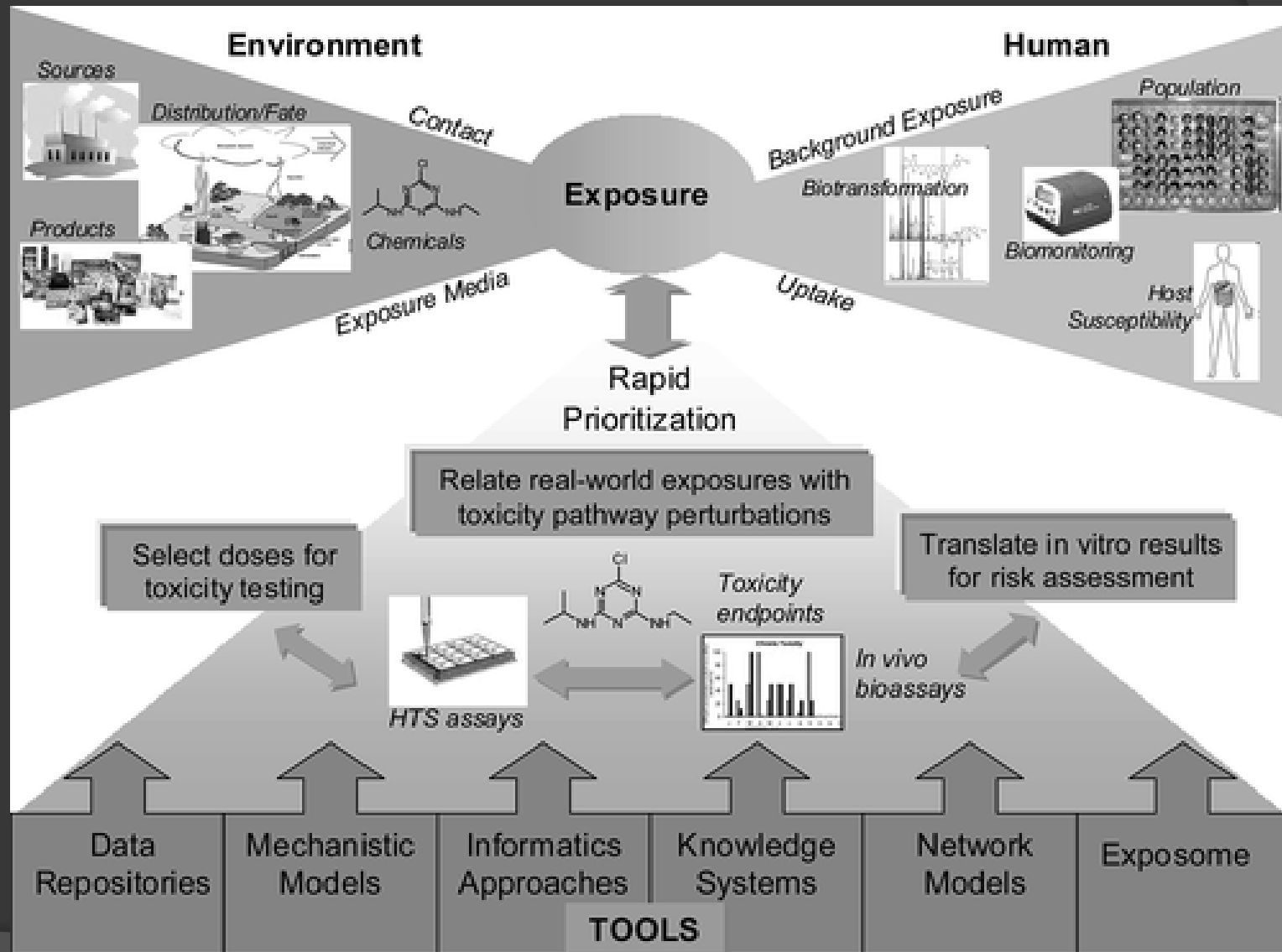
Network /
Connectivity

0: No interoperability

Increasing Capability



Disciplinary Interoperability



US EPA ORD Pathfinder Innovation Project: Global to Genome (G2G)

- GOAL: *Specification* of a Computational Platform for Agency-wide, Seamless Data Flow and Computational Modeling in Support of Health, Ecological, and Climate Risk Characterizations

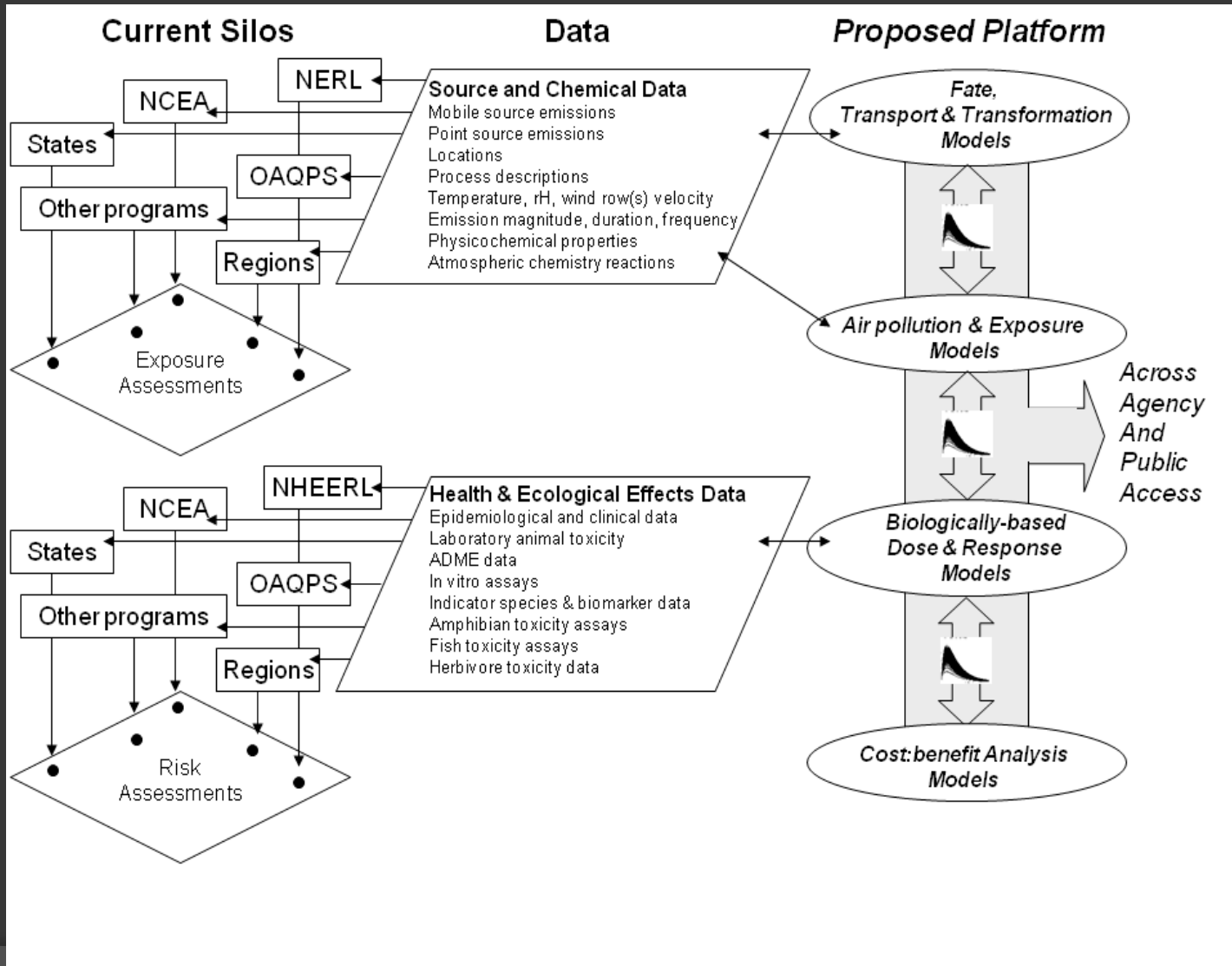


G2G Team Members

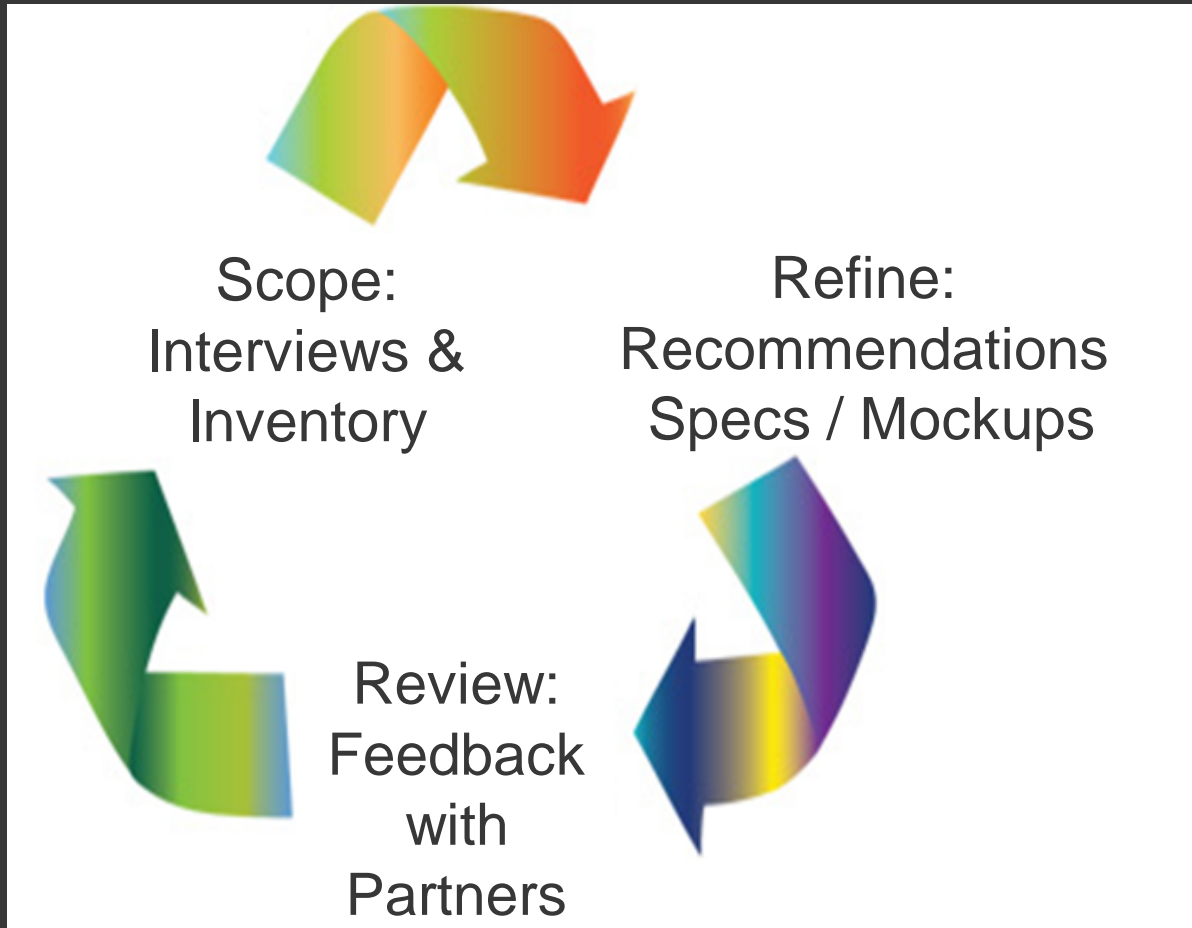
- NCEA: *Annie Jarabek*,
David Miller, and Lyle Burgoon
- NHEERL: *Rory Conolly*,
Stephen Edwards and William LeFew
- NERL: Gene Whelan, Cecilia Tan and
Michael Breen
- OSIM: David Lyons



G2G: Problem and Promise



G2G Process



- ⦿ Internal
 - ORD labs and centers
 - Program Offices
 - Regions
- ⦿ External vetting

SOT CCT Workshop

Building for Better Decisions (B4BD): Multi-scale Integration of Human Health and Environmental Data

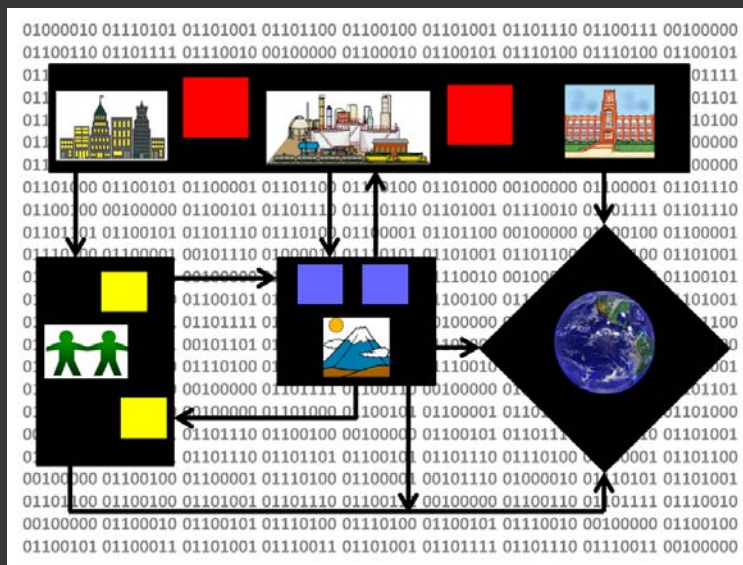
- Held @ EPA RTP campus May 8 – 11, 2012
- https://www.toxicology.org/ai/meet/cct_b4bd.asp
- International workshop with sponsors across sectors

Professional		
Government	Society	Private Sector
US ACE	SETAC	American Chemistry Council
USDA	SRA	Environ International Corp.
US FDA	ISES	ICF International
USGS	iEMSs	OGC
US NRC		OpenMI
PNNL		TERA

SOT CCT B4BD

Organizing Committee

- ⦿ Lyle Burgoon, Co-chair, US EPA NCEA RTP
- ⦿ Robinan Gentry, Co-chair, Environ Corporation
- ⦿ Annie Jarabek, Co-chair, US EPA NCEA IO
- ⦿ Richard Corley, US DOE PNNL
- ⦿ George Daston, Proctor and Gamble, Co., Inc.
- ⦿ Paul Price, The Dow Chemical Co.
- ⦿ Edward Perkins, US Army Corps of Engineers
- ⦿ Glenn Suter, US EPA NCEA Cincinnati
- ⦿ Bruce Vigon, SETAC
- ⦿ Gene Whelan, US EPA NERL Athens
- ⦿ Timothy Zacharewski, Michigan State University




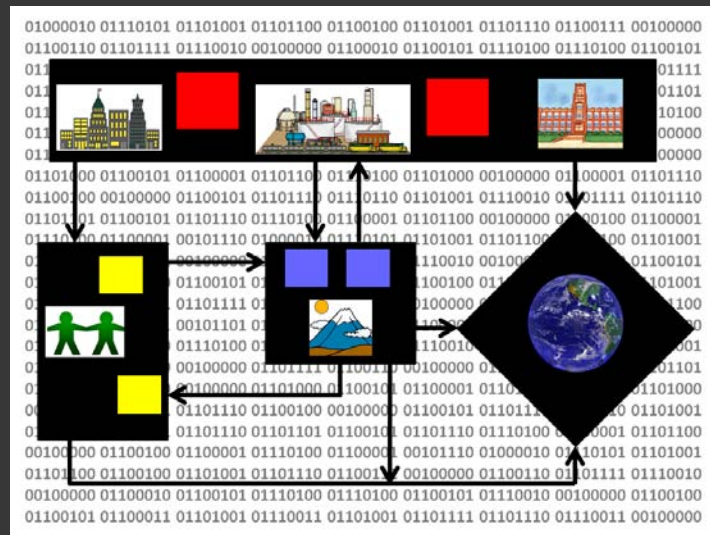
B4BD Context: Motivation, Concepts and Challenges

- Protecting the public health and environment requires analysis, translation, and integration of data along source to effect pathways
- Technological advances challenge scientific community to harness larger and increasingly complex data in a transparent, integrated fashion

Challenges



- Inadequate scientific understanding, especially of “narratives”
- New types of data 
- User expectations
- Limitations / vulnerability of technologies
- Proprietary considerations
- Regulatory process and institutional structures
- Doing more with less



Process

- ⦿ Plenary presentations
 - Across sectors and disciplines
 - Emphasis on IT
- ⦿ Theme discipline discussions
 - Invited participant experts across sectors
 - “Ambassadors” from other themes
- ⦿ Reports from themes to entire plenary
- ⦿ Attendees across all
- ⦿ Poster session / reception by themes

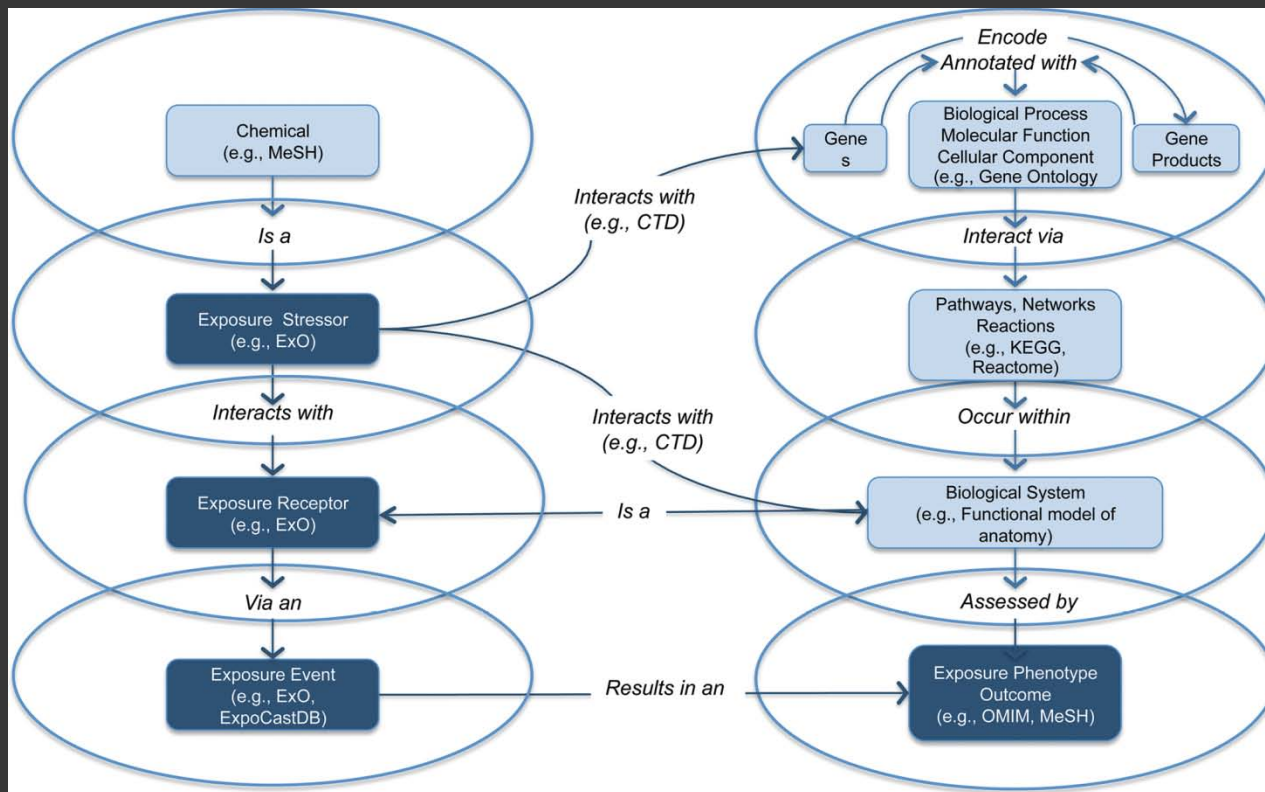
SOT CCT B4BD

Disciplinary Themes

- **Theme A:** Exposure Assessment, Transport and Transformation
 - Paul Price, The Dow Chemical Co. / Brenda Barry, American Chemistry Council
- **Theme B:** Ecotoxicological Assessment, Ecosystem Services, Climate Change
 - John Johnston, US EPA NERL Athens / Glenn Suter, US EPA NCEA Cin
- **Theme C:** Dose-response, Tox21 and Risk Assessment
 - Michael Waters, Integrated Laboratory Systems, Inc. / Rory Conolly, US EPA NHEERL and Lynne Haber, TERA
- **Theme D:** Life-cycle / multi-criteria Assessment and Cost:Benefit Analysis
 - Bruce Vigon, SETAC / Christina Powers, US EPA NCEA
- **Theme E:** Information Technology
 - Roger Perkins, US FDA NCTR / Bernadette Highland, 3 Round Stones

Broader Biological Context for Exposure

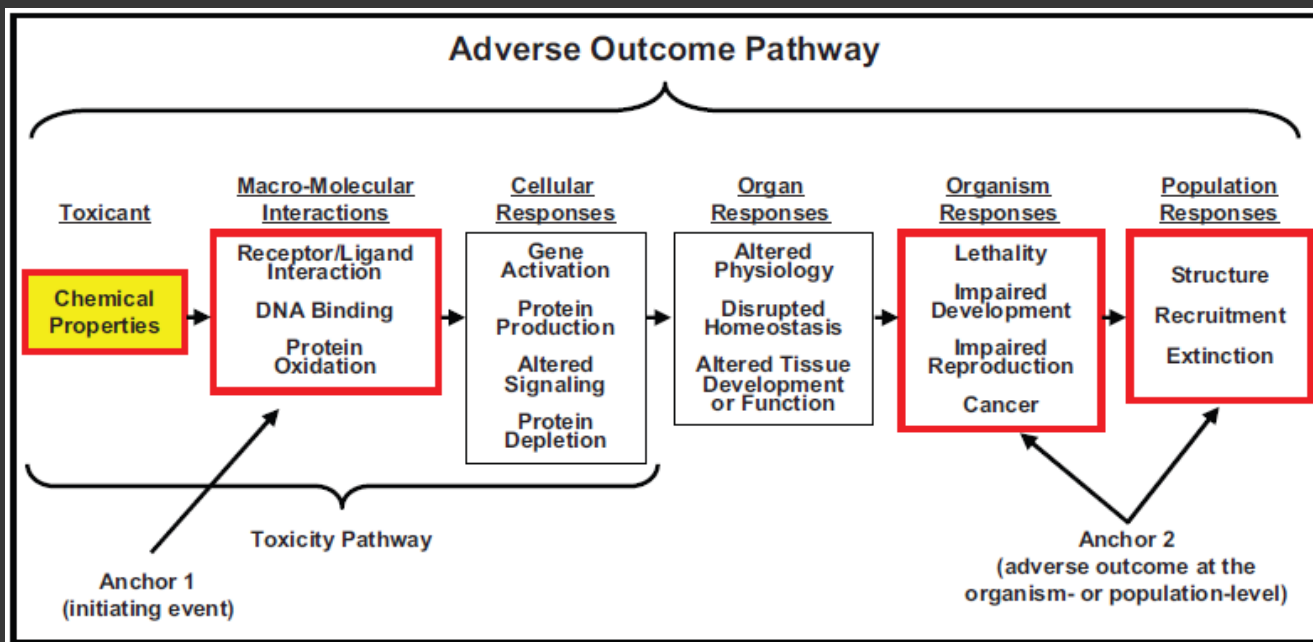
Theme A



Mattingly et al (2012) Environ. Sci. Technol. In press.

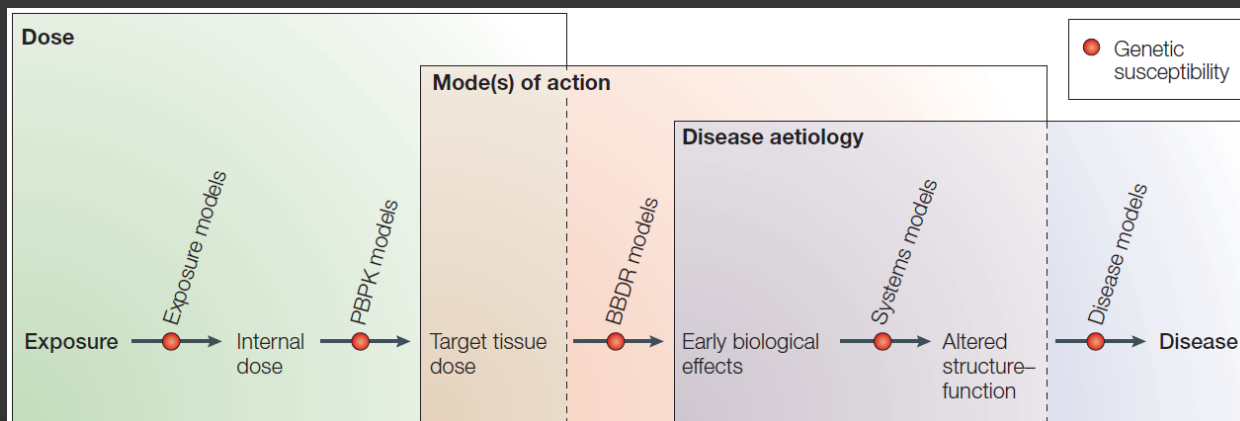
Adverse Outcome Pathways

Theme
B



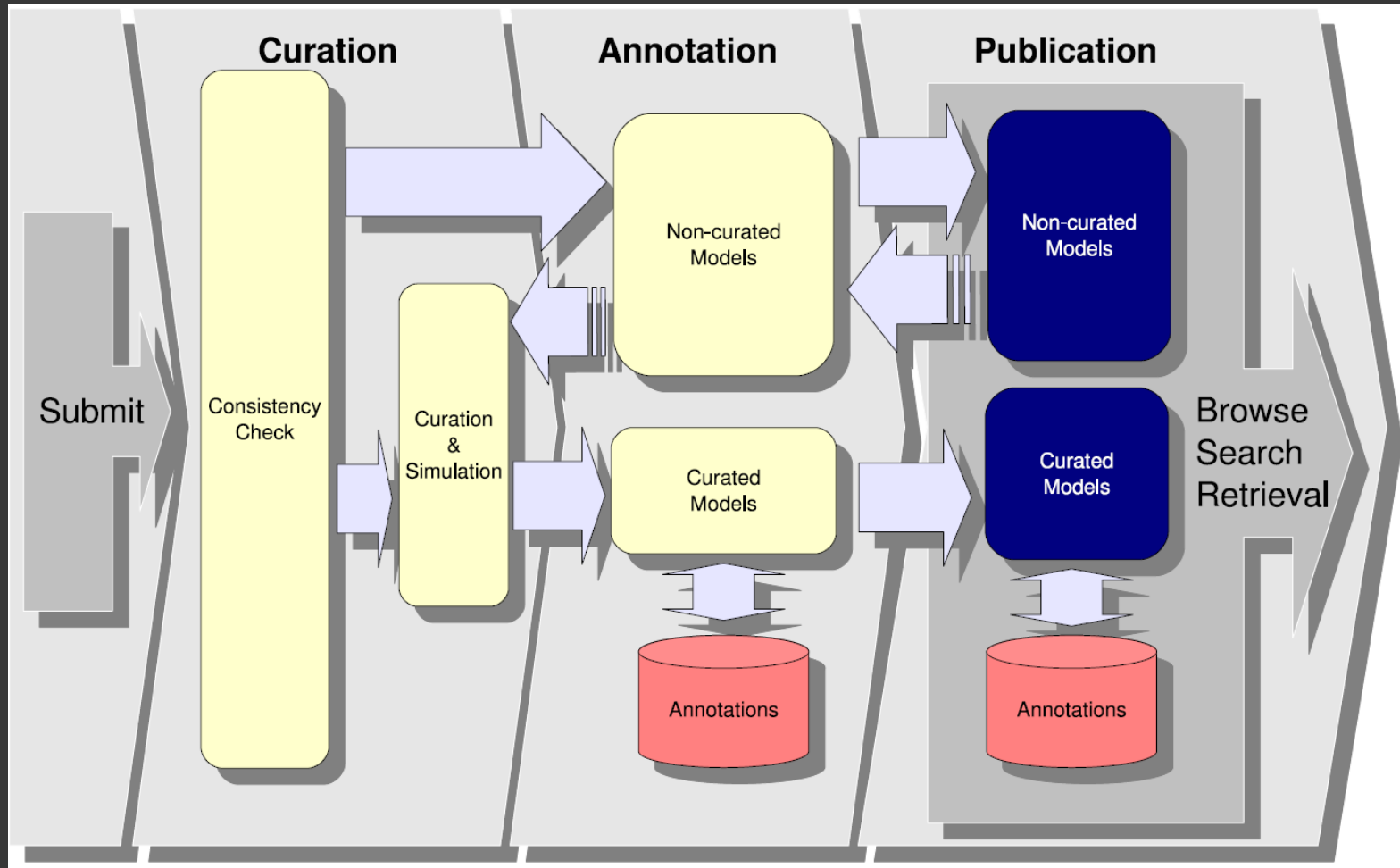
Ankley et al. (2010) Environ. Toxicol. Chem. 29, 730.

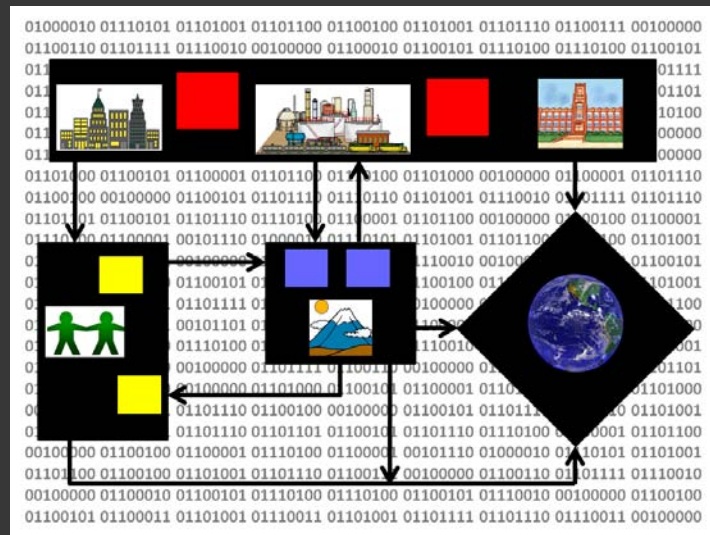
Theme
C



Waters and Fostel (2004) Nature Rev. Genetics, 5, 936.

Data Management Issues





Products

- ◎ Disciplinary manuscripts
 - State of the science / best practice
 - Identify IT needs to advance on path forward
- ◎ Synthesis summary of meeting
- ◎ Glossary (e.g., vein)
- ◎ Community of Practice of Integrated Environmental Modeling (CIEM) survey
- ◎ Demonstration project

Benefits

- ⦿ Transparency, documentation, communication
 - Tractable work flows / calculations
 - Assumptions and limitations
 - Propagation of uncertainty
- ⦿ Efficiency
 - Dockets for regulatory processes / reviews
 - Real-time response to peer review
 - Comparison of alternative approaches
 - Template models and resource leveraging
- ⦿ Scalable
 - Fit-for-purpose applications
 - Repurposing of data
- ⦿ Coordination and consistency across Agency
 - Education
 - Contemporary methods and approaches

Recommendations

- VISION: Empower *BOTH* research and decision making
- DEFINITION: Address *interoperability* in 2 domains
 - Technical: RDF, software, machine code, etc.
 - Cultural: Community of practice, processes
- NEEDS
 - ACCESS to discover, collect, and integrate data in a coordinated fashion to inform research and decision making
 - REPURPOSE data for own analysis REQUIRES context for data (meta data) including annotation and curation history
- BENEFITS
 - Increased integration of data that could be truly iterative
 - Increased transparency, efficiency, and communication
- IMPLEMENTATION
 - Standards and recommendations
 - Demonstration projects



Proposed G2G Platform for EPA

- ◎ Research and Decision Support Tool
 - CENTRAL on-line Agency location containing or referencing databases, models, algorithms, spreadsheets, etc.
 - An information / metadata hub containing current research, desired extensions / needs, regulatory policy positions, etc.
- ◎ Workbench
 - An online tool to find and show connections between models so that inputs, outputs, and uncertainty are explicitly and transparently propagated. Provide a portal to platforms that facilitate this process (e.g. FRAMES).



G2G Specifications: Technical

- ◎ Research and Regulatory Support Tool
 - Access is primary issue (CBI? Public? Contractors?)
 - Simplicity of data entry and user interface
 - Repurposing of data, models
 - Thorough documentation via metadata
 - Transparency of assumptions and uncertainty
 - Encourage collaborative development
- ◎ Workbench
 - Emphasize standards for interoperability
 - Enable and assist in construction of connectors
 - Simplify comparison of models
- ◎ Data and Model Management
 - Support migration of data and models (e.g., PI based)
 - Archiving, documentation, decision records, policy updates



G2G Specifications: Cultural

◎ Realization

- Top-level champion is essential
- Requires sufficient management support and resources

◎ Implementation

- Institutional home is necessary
- Do not create a behemoth – leverage extant systems
- Team to develop, disseminate and coordinate
- Position tool to be useful to both research and regulatory support

◎ Adoption

- Incentives for sharing models and publication of data
- Content must be maintained and updated
- Support and resources from development AND for duration
- Utilize programmatic and disciplinary champions



G2G Mockup: User entry

Browser address bar: <http://g2g.epa.gov>

EPA Global to Genome

Welcome Annie! [sign out](#)

Home Tools Data Workbench

Like 64 people like G2G. Be the first of your friends. [advanced search](#)

Notifications

[change notifications settings](#) ?

- 1** [Kevin Fischer](#) rated and left a comment on [Simple Tool](#)
Tool created in R
43 minutes ago
- 2** [Stephen Edwards](#) downloaded of [Simple Tool](#)
Tool created in R
6 hours ago
- [Stephen Edwards](#) requested download of [Some Chemical Model I Created](#)
Tool created in R
6 hours ago
- [Michael Breen](#) joined [G2G](#) project
Collaboration Team Update
43 minutes ago [6 members](#)

Recommended

- 3** [Jane Doe](#) uploaded a spreadsheet, [Kin locations of Wisconsin, Illinois, & Ohio](#)
Short description of the spreadsheet
19 minutes ago
- [Cecilia Tan](#) uploaded a csv, [air_samples.csv](#)
Short description of the csv
yesterday
- [Jane Doe](#) registered a database, ["ToxBase"](#)
Short description of the database
Jan 12, 2012

Economic **Social** **Environmental**

Source F&T Exposure Dose Effects Outcome Cost Assess Benefit

89 876 57

PBPK

- Deltamethrin Rat/Human PBPK
- PBPK Human Styrene Exposure
- PBPK Mixture Lumping Algorithm
- Permethrin Human PBPK
- Permethrin Rat PBPK

Date

Relevance Popularity Rating

wb simplify mixture systems about which parameters are known in yesterday (23) ★★★★★

wb [Permethrin Rat PBPK](#) Short description of this model 2 days ago (23) ★★★★★

wb [Permethrin Human PBPK](#) Short description of this model yesterday (23) ★★★★★

wb [ToxBase](#) integrated Tox-type stuff database Short description of the database yesterday (23) ★★★★★

4

5



G2G Mockup: Tool Page

Browser: G2G
Address: http://g2g.epa.gov

EPA Global to Genome

1 Permethrin Rat PBPK

Rat physiologically based pharmacokinetic model for permethrin

Home Tools Data Workbench

Annie's G2G Sign Out advanced search

60% tool metadata complete

2 model metadata (What is this?)

(23) ★★★★★ [Contribute](#)

3 Purpose A rat PBPK model was developed to make use of in vitro and in silico information pertinent to permethrin pharmacokinetics. The rat model is used to evaluate the in vitro-in vivo extrapolation and to better understand the sources of uncertainty that may impact a human model of permethrin disposition.

Author(s) [Rogelio Tornero-Velez](#), [Jimena Davis](#), [Edward Scollon](#), [James Starr](#), [Woodrew Setzer](#), [Michael Goldsmith](#), [Daniel Chang](#), [Jianping Xue](#), [Valerie Zartarian](#), [Michael DeVito](#), [Michael Hughes](#)

Version v 1.1.2 [explore version history](#)

Citation PubMed [PMID: 22859315](#)

Software MatLab

Peer Review [Published Manuscript](#), 2010

Domain **economic** social environmental

Sub-domain Source F&T Exposure Dose Effects Outcome Cost Benefit

Status [Endorsed by some program office of US EPA.](#)

Input(s)

Output(s)

4

Permethrin Rat PBPK
perm_rat-1_1_2.zip

[add to workbench](#) 22 downloads 34 comments

No known linkage

2 known inheritance
Parent of [Permethrin Human PBPK](#)
Child of [Deltamethrin Rat/Human PBPK](#)

5

Comments

[Add a comment](#)

[Mary Smith](#) -Considering using this as a basis for a Cypermethrin Human PBPK
43 minutes



G2G Mockup: Workbench

G2G

http://g2g.epa.gov

EPA Global to Genome

Home Tools Data Workbench

Permethrin PBPK | Deltamethrin Rat/Human PBPK | Untitled Workbench 01 | Annie's G2G | Sign Out | advanced search

Show Inheritance Hide Connection Workbench Add Input Add Output Add Model or Data Search

Workbench 001

D4EM
137 Surface Impoundments

HE2RMES
Inhalation

Exposure Estimates
Point Estimate *type point*
Exposure *type decimal*

Exposure Estimates
Census Tract *type polygon*
Exposure *type integer*

Workbench 002

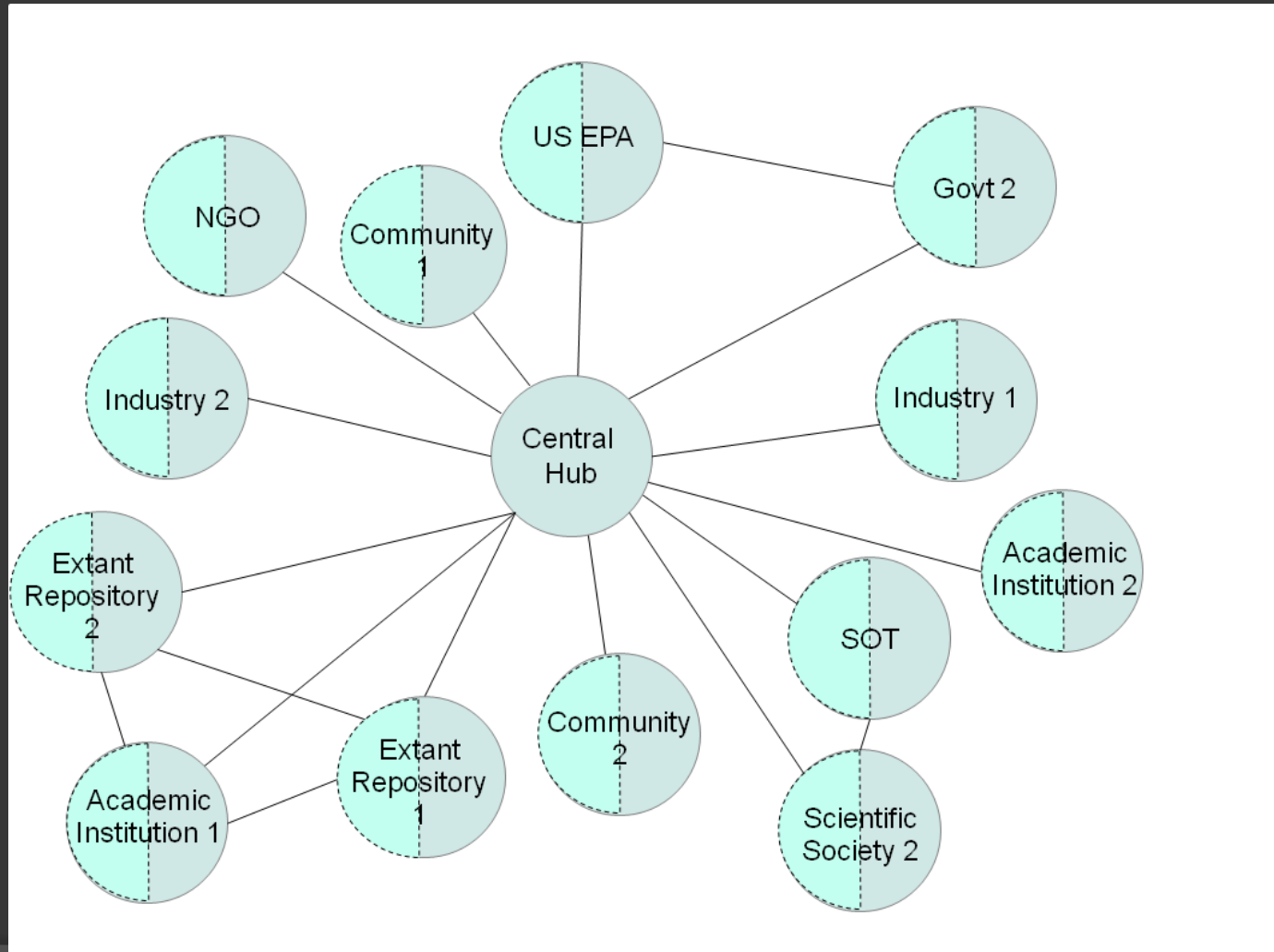
NATA
National Scale Air Toxics Assessment

ASPEN
Assessment System for...

+ Zoom to fit -



G2G in the Future: Federation



SOT CCT B4BD Demonstration Project

- ⦿ Engage all disciplines
 - LCA frame for problem
 - Embedded computational/scientific issue for each theme
- ⦿ Candidates under consideration
 - Mercury in light bulbs
 - Arsenic in different drinking water sources
 - BPA
- ⦿ Others?
- ⦿ Contact: jarabek.annie@epa.gov