

For Conference Purposes Only



Improving Transparency Among Scientists: The Role of Shared Terminology in Multidisciplinary Science

*Lynne Petterson, Ph.D., ORD/OSIM/IMSB
Gail Hodge, Information International Assoc.
OEI National Symposium
May 12, 2010*



National Challenges are Multi-disciplinary

- Research is becoming more multi-disciplinary and interdisciplinary
- Different disciplines have different vocabularies
 - Same term with different definitions
 - Different terms for the same concept
- Productive collaborations have agreed upon, shared language and meaning, avoiding misunderstanding and rework



Managed Vocabulary Goals

- Collect EPA ORD terms and multiple definitions
- Mediate terms and meanings
- Applications
 - Improved search capability, (e.g., CAFO *and* confined animal feeding operation/ sludge *and* biosolid)
 - Automated metadata creation
 - Classify non-text objects such as data/sets
 - Link systems together
 - Variable name in one data/set is called x, and in another data/set is called y, but means SAME thing, vocabulary helps provide linkage

Scope of Managed Vocabulary

- 18 high level categories show the broad scope of EPA and ORD's research
- Range from biology and technology to society and economics (sustainability)

ORD Facets

Audience

- Biological Substance
- Built Environment
- Chemical Substance

Content Type

- Ecosystem
- Environmental Event
- Environmental Technology

EPA Operations

Field of Study

Laws, Regulations & Treaties (selected)

- Location
- Measurement
- Natural Environment
- Organization and Industry
- Outcome
- Research
- Society and Economy

Managed Vocabulary Status

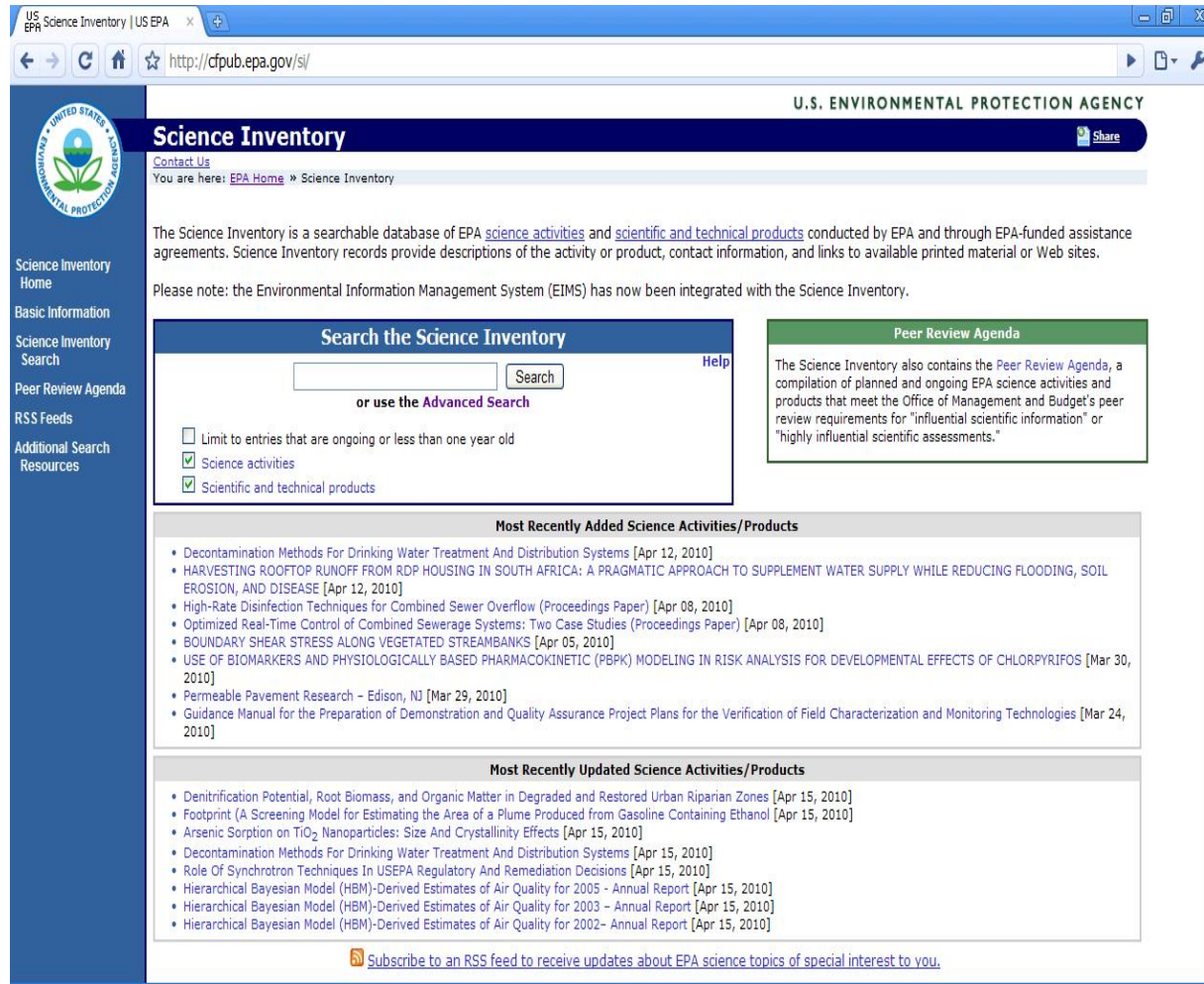
- **Collected Terms**
 - Interviews
 - Key documents and web sites
 - Other EPA glossaries from Terminology Services
 - Other vocabularies, especially NIH, USGS and Agriculture
- **Organized terms in a framework**
 - Developed a framework structure based on EPA Mission
 - Placed terms in the framework categories
 - Created term hierarchies and relationships
 - Currently >5,800 terms; >3,500 relationships of which 225 are to acronyms or abbreviations
 - Began collecting definitions
- **Review**
 - Engaging SMEs
 - This is how you can help

Use Cases of Different Types

- Developing use cases to help identify additional vocabulary, test if structure will support potential uses, and identify web services, governance, etc.
- System Oriented
 - Science Inventory
 - Integrated Modeling
- Human Oriented
 - MyProject Glossary

Science Inventory

- Improve search and access to ORD research, including data
- Supplement current use of uncontrolled keywords



The screenshot shows the EPA Science Inventory website interface. At the top, the browser address bar displays "http://cfpub.epa.gov/si/". The page header includes the EPA logo and the text "U.S. ENVIRONMENTAL PROTECTION AGENCY". The main content area is titled "Science Inventory" and includes a "Contact Us" link and a breadcrumb trail: "You are here: EPA Home » Science Inventory".

The page describes the Science Inventory as a searchable database of EPA science activities and scientific and technical products. It notes that the Environmental Information Management System (EIMS) has been integrated with the Science Inventory.

A search section titled "Search the Science Inventory" features a search input field, a "Search" button, and a "Help" link. Below the search field, it says "or use the Advanced Search". There are two checkboxes: "Limit to entries that are ongoing or less than one year old" (unchecked) and "Science activities" (checked). A third checkbox, "Scientific and technical products", is also checked.

Two sections are highlighted with green boxes: "Peer Review Agenda" and "Most Recently Added Science Activities/Products". The "Peer Review Agenda" section states that the Science Inventory also contains the Peer Review Agenda, a compilation of planned and ongoing EPA science activities and products that meet the Office of Management and Budget's peer review requirements for "influential scientific information" or "highly influential scientific assessments."

The "Most Recently Added Science Activities/Products" section lists several items, including:

- Decontamination Methods For Drinking Water Treatment And Distribution Systems [Apr 12, 2010]
- HARVESTING ROOFTOP RUNOFF FROM RDP HOUSING IN SOUTH AFRICA: A PRAGMATIC APPROACH TO SUPPLEMENT WATER SUPPLY WHILE REDUCING FLOODING, SOIL EROSION, AND DISEASE [Apr 12, 2010]
- High-Rate Disinfection Techniques for Combined Sewer Overflow (Proceedings Paper) [Apr 08, 2010]
- Optimized Real-Time Control of Combined Sewerage Systems: Two Case Studies (Proceedings Paper) [Apr 08, 2010]
- BOUNDARY SHEAR STRESS ALONG VEGETATED STREAMBANKS [Apr 05, 2010]
- USE OF BIOMARKERS AND PHYSIOLOGICALLY BASED PHARMACOKINETIC (PBPK) MODELING IN RISK ANALYSIS FOR DEVELOPMENTAL EFFECTS OF CHLORPYRIFOS [Mar 30, 2010]
- Permeable Pavement Research - Edison, NJ [Mar 29, 2010]
- Guidance Manual for the Preparation of Demonstration and Quality Assurance Project Plans for the Verification of Field Characterization and Monitoring Technologies [Mar 24, 2010]

The "Most Recently Updated Science Activities/Products" section lists several items, including:

- Denitrification Potential, Root Biomass, and Organic Matter in Degraded and Restored Urban Riparian Zones [Apr 15, 2010]
- Footprint (A Screening Model for Estimating the Area of a Plume Produced from Gasoline Containing Ethanol) [Apr 15, 2010]
- Arsenic Sorption on TiO₂ Nanoparticles: Size And Crystallinity Effects [Apr 15, 2010]
- Decontamination Methods For Drinking Water Treatment And Distribution Systems [Apr 15, 2010]
- Role Of Synchrotron Techniques In USEPA Regulatory And Remediation Decisions [Apr 15, 2010]
- Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2005 - Annual Report [Apr 15, 2010]
- Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2003 - Annual Report [Apr 15, 2010]
- Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2002- Annual Report [Apr 15, 2010]

At the bottom of the page, there is a link to "Subscribe to an RSS feed to receive updates about EPA science topics of special interest to you."

Integrated Modeling

- Develop and utilize a standard, common nomenclature to support the infrastructure for integrated modeling
 - Describe the models which vary by discipline and media
 - Common vocabulary for practice of modeling
- Facilitate quality assured communication of data and knowledge

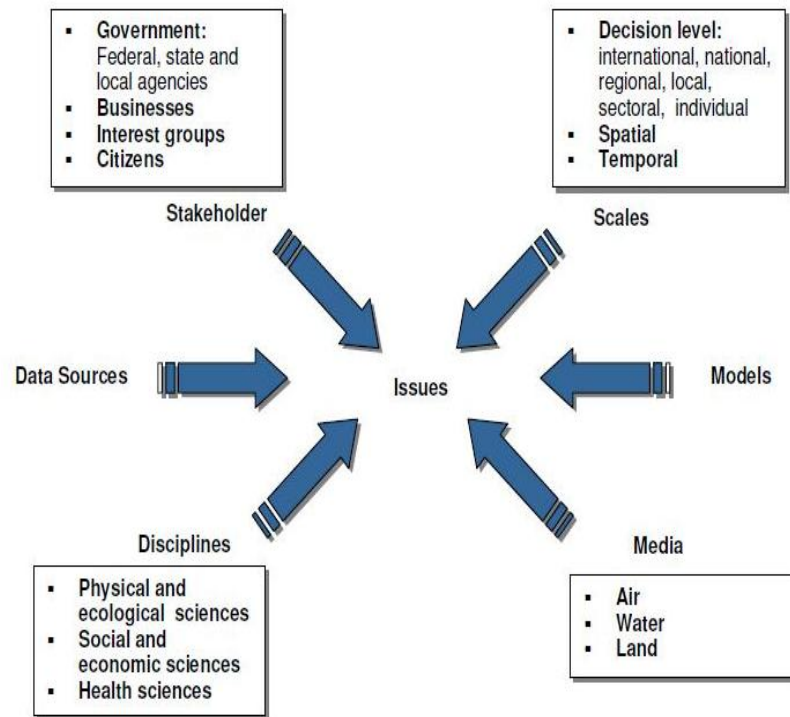


Figure 3: Issues of Scale and Dimensions of Integrated Modeling¹⁷

CREM Integrated Modeling White Paper, EPA 100/R-08/010

Proposed MyProject Glossary Functions

- Set up project glossary and permissions
- Search other glossaries for relevant terms
- Copy, add and modify terms and definitions
- Add acronyms, abbreviations and synonyms
- Gain consensus on terms and definitions for this project
- Stamp terms and definitions as having been used by this project

Set up project
glossary and
permissions

Select Content

- Search other glossaries for relevant terms
- Copy, add and modify terms

Develop Content

- Develop/modify terms and definitions
- Add acronyms, abbreviations and definitions

Agree on and Stamp
terms and definitions

MyProject Glossary Concept - Main Page

Estuary Project Glossary - ... x

https://wiki.epa.gov/terminology/index.php/Estuary_Project_Glossary

DTurner my talk my preferences my watchlist my contributions log out

page discussion edit history delete move unprotect watch

Estuary Project Glossary

Contents [hide]

- 1 Welcome
- 2 Import
- 3 Complete List of Terms
- 4 Export

Navigation

- Main Page
- Help

terminology

- System Lifecycle Management Vocabulary
- Chesapeake Bay Prog. Vocabulary

search

Go Search

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Welcome [edit]

This wiki has been created to support your project's efforts to develop a common vocabulary. The [MyProject Glossary](#) allows the import of whole glossaries or selected terms from Terminology Services, the repository for EPA glossaries. The wiki interface promotes collaboration while your group discusses and edits the terms and definitions. At the end of the process the terms can be exported to Terminology Services.

Import [edit]

Import Glossary or Terms

Complete List of Terms [edit]

Algae
Algae Blooms

Export [edit]

Export Term Pages

MyProject Glossary Concept – Discussion Page


Talk:Algae - Terminology S... x


https://wiki.epa.gov/terminology/index.php/Talk:Algae

DTurner my talk my preferences my watchlist my contributions log out

page **discussion** delete protect




Talk:Algae

[View archived threads](#) 

Sorting order: 

Remember this preference

Algae (1)

[History](#)  [Move](#)  [Watch](#) 

Term: Algae

Vocabulary: Estuary Project Glossary

Def1: Simple rootless plants that grow in sunlit waters in proportion to the amount of available nutrients. [Terms of Env]

Def2:

Source Note:







Scope Note:

Editorial Note:






Acronym:


Preferred Term:


Related Term: Algae Blooms

DTurner (Talk | contribs | block) Edited  16:08, 20 April 2010 [Edit](#)  [History](#)  [Permalink](#)  [Delete](#)  [Reply](#) 

Def2: Simple one-celled or many celled plants capable of photosynthesis, usually aquatic. [OWOW Glossary]

DTurner (Talk | contribs | block) 16:10, 20 April 2010 [Edit](#)  [History](#)  [Permalink](#)  [Delete](#)  [Reply](#) 

[Create a new template](#)  (System Administrator Use Only)



navigation

- Main Page
- Help

terminology

- System Lifecycle Management Vocabulary
- Chesapeake Bay Prog. Vocabulary

search

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version

MyProject Glossary Concept – Term Page

Algae - Terminology Services x

https://wiki.epa.gov/terminology/index.php/Algae

DTurner my talk my preferences my watchlist my contributions log out

page discussion edit history delete move unprotect watch

Algae

Term: Algae

Vocabulary: Estuary Project Glossary

Def1: Simple rootless plants that grow in sunlit waters in proportion to the amount of available nutrients. [Terms of Env]

Def2: Simple one-celled or many celled plants capable of photosynthesis, usually aquatic. [OWOW Glossary]

Source Note:

Scope Note:

Editorial Note:

Acronym:

Preferred Term:

Related Term: Algae Blooms

[Export Term](#)

[Return to Estuary Project Glossary](#)

navigation

- Main Page
- Help


terminology

- System Lifecycle Management Vocabulary
- Chesapeake Bay Prog. Vocabulary

search

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link



Looking Forward: The Semantic Web

- Managed Vocabulary can be used to support semantic mediation
- Relationships developed between terms can be expressed in Web 3.0 formats such as RDF to support approaches such as Linked Data which will be discussed in more detail in a later presentation
- Through MyProjects begin to build the semantic knowledge base that is needed to improve the understanding, use and management of EPA science products



Contacts

Lynne Petterson, Ph.D.

Office of Science Information
Management/
Information Management Support
Branch
Petterson.Lynne@epa.gov
919-541-3582

Gail Hodge

Advanced Informatics Group
Information International Assoc.,
Inc. (IIa)
ghodge@iiaweb.com
865-742-5430