

# Informatics Approaches for the Reuse and Modeling of Heterogeneous Mercury Data

**MELINDA NEVILLE**

Dept. Ecol & Env Sci

**KATE BEARD**

School of Comp & Info Sci

**UNIVERSITY OF MAINE**



Snow sampling 3-5-04  
Site: 45  
Initials: TA  
Depth: 19.5

Wt. empty tube: 3.37  
Depth of snow: 23  
Length of core: 23  
Wt. tube + core: 4.10  
Remarks: down to litter. fir  
denser lower Southside.

Wt. empty tube: 3.35  
Depth of snow: ~~25~~ 19.5  
Length of core: 19.5  
Wt. tube + core: 3.96  
Remarks: had a hard time getting the  
core because of rocks

Snow sample collection  
Hg samp remarks  
lab wts  
4.31

sample collection  
Hg samp remarks touch ORANGE  
ceder branch  
sample pulled out. ice block in  
on top.

4.84  
lab wt = 4.84 kg!  
- Double bag -

## “...Informatics?”

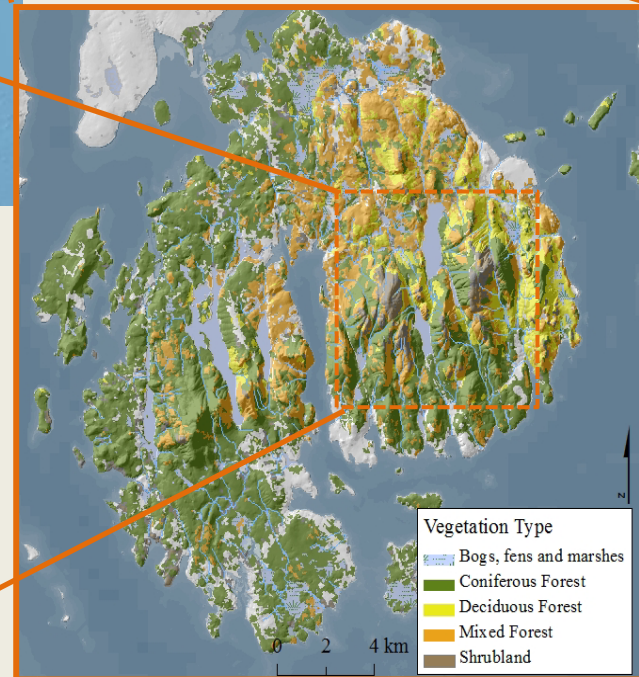
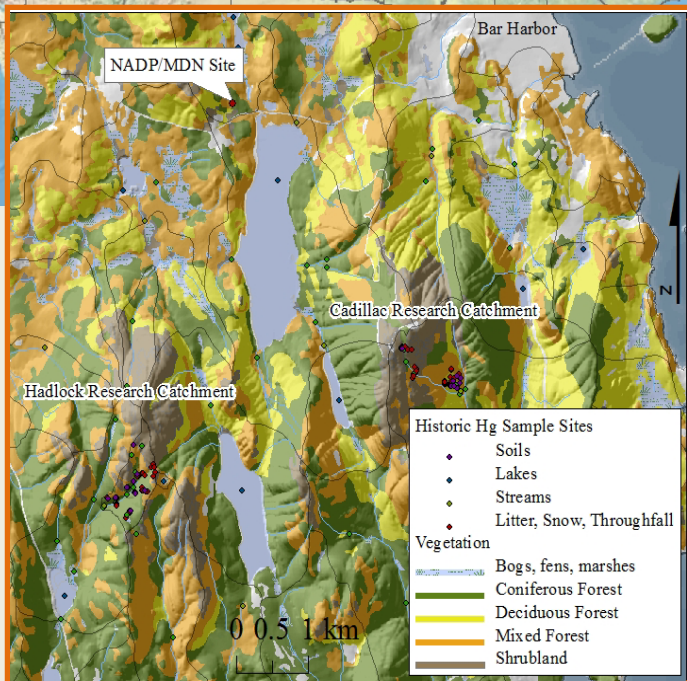
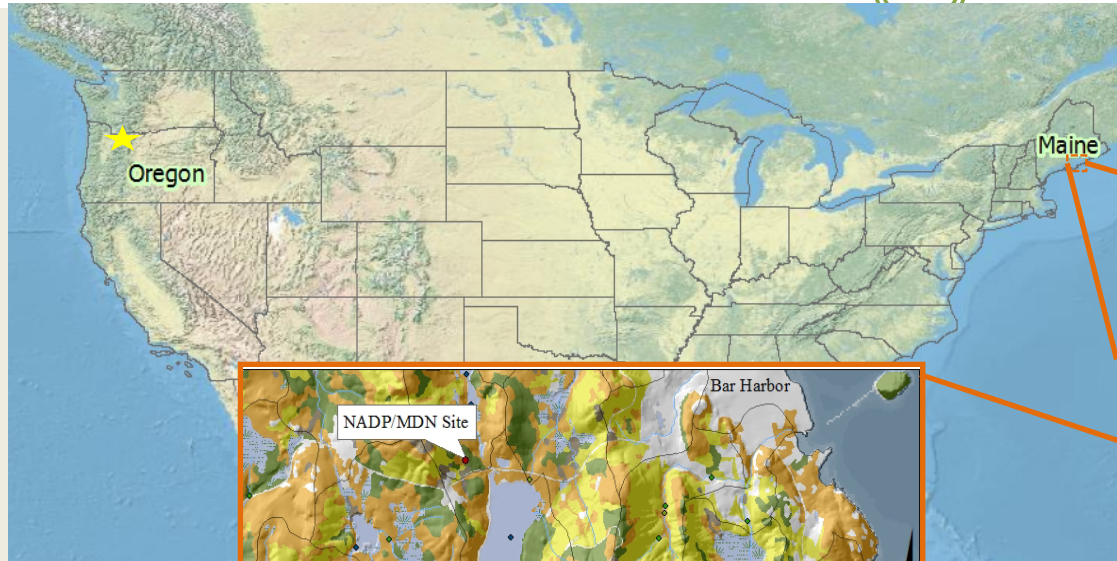


- Coupling information with **computing technology**
- Data mining & **information analysis**
- Creating the **cyberinfrastructure** for a knowledge economy (Atkins, 2003)
- The application of information science to facilitate scientific research goals

Research Goal:

**Reuse and Modeling of  
Heterogeneous Mercury Data**

# Acadia National Park



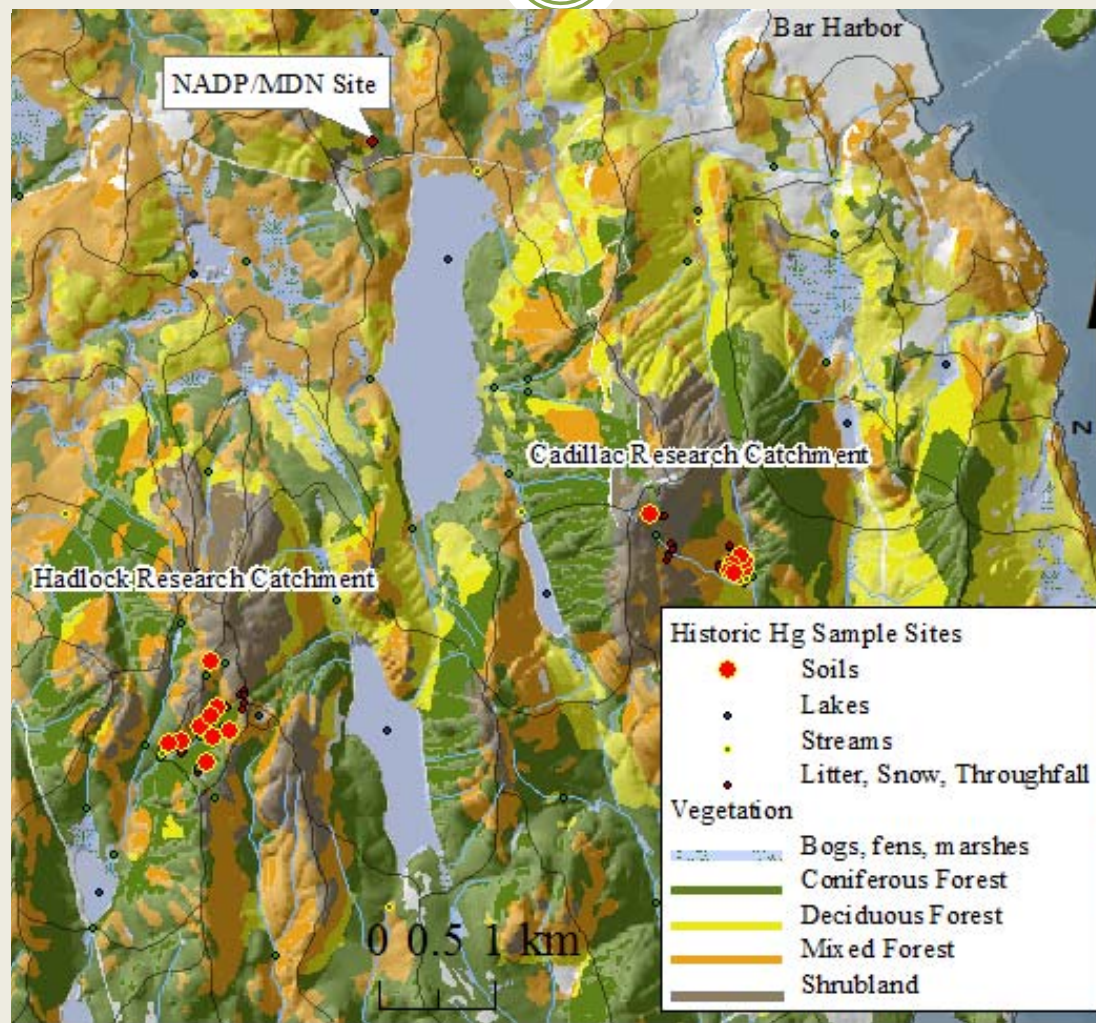
# Legacy Hg studies at Acadia National Park



<b>Data source</b>	<b>Media</b>	<b>Sampling Method</b>	<b>Time Span</b>	<b>Samples</b>
NADP MDN	Wet precipitation	Aerochem Metrics	1996-	700
EPA	Sediment	Unknown	2000 - 2002	34
EPA	lobster		2000 - 2002	35
Gulf Watch	mussels	Unknown	1993 - 2001	200
Peckenham	Streams	grab samples	1999 - 2004	150
Sheehan	foilage	baskets	2003 - 2004	153
Johnson	wet+dry precip.	funnel collectors	1999 - 2004	500
Kahl & Nelson	streams	grab samples	1999 - 2003	107
Nelson	snow	event/season	2004 - 2005	80
Burgess	fish, plankton, inverts		1997	200
Bank	amphibians		2003	200
Longcore	birds	eggs, food, bodies	1997 - 1999	35
Ruck	soils	excavation	1999	40
BRI	mink, otter	tissue, fur	2001	3
BRI	loons	blood, feathers, eggs	1997 - 2002	11
Welch	eagles	blood, feathers	1991 - 1992	7

Bank et al 2007; Johnson et al 2007; Kahl et al, 2007; Longcore et al 2007; Nelson et al 2008; Ruck 2002; Peckenham et al 2007; Sheehan et al 2005

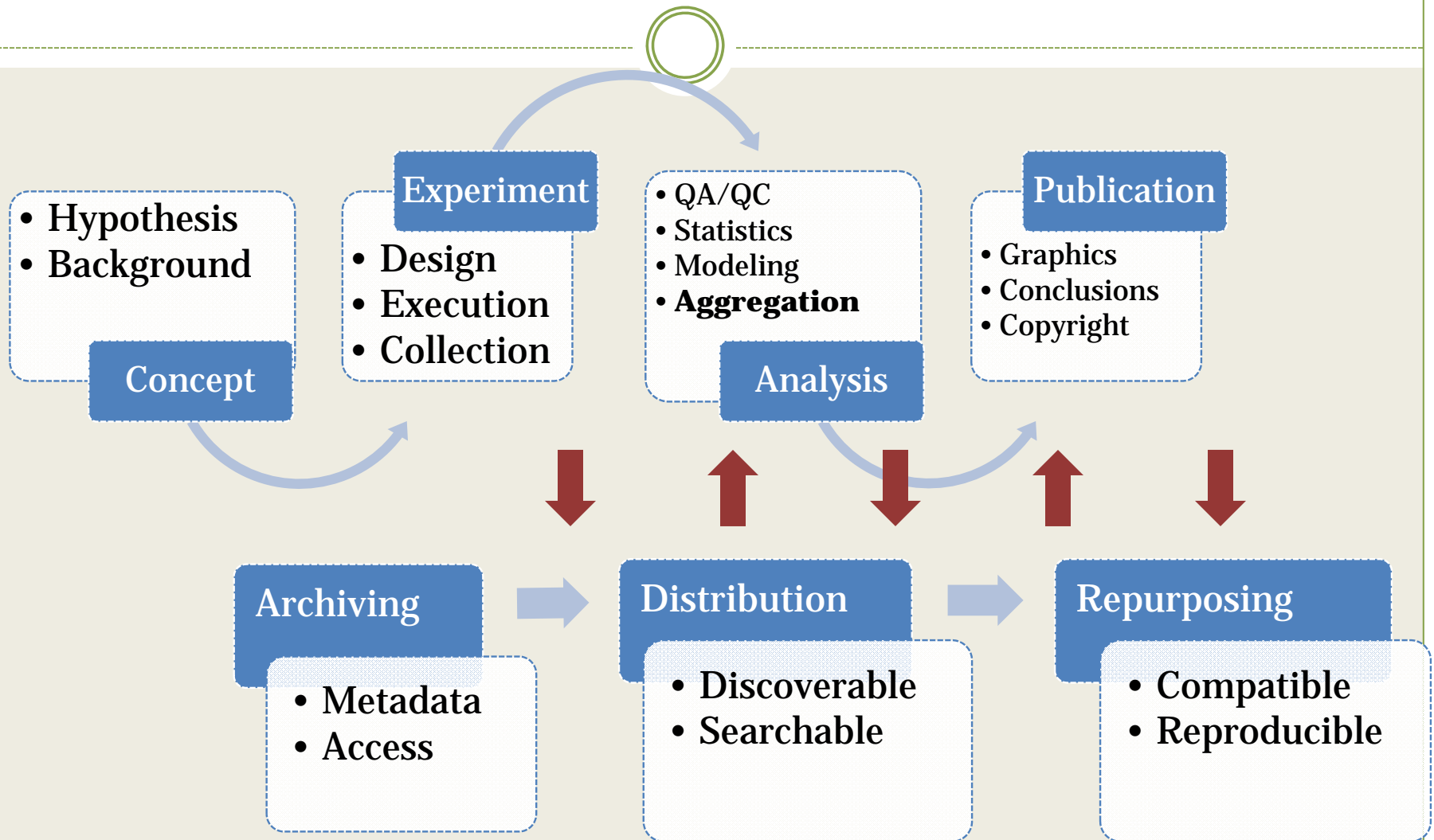
# Acadia National Park



# Research Questions

- Do current metadata standards support comparison across datasets?
- What approaches should be adopted to enhance data repurposing?
- Can the semantics of Hg research be standardized sufficiently to allow reasoning across large collections of disparate data?

# When do we think about repurposing?



Snow sampling <sup>left core 2 @ C3</sup> Date: 3-5-04

Site: C3A  
Initials: MA  
TH

Time: 11:45  
Air temp: 0°

Depth survey: BLUE

Wt. empty tube: 3.37

Depth of snow: 23

Length of core: 23

Wt. tube + core: 4.10

Remarks: down to litter. fluffy in low  
denser lower Southside.

Snow sample collection

Hg samp remarks

lab wt:  
4.31

Non-Hg samp remarks

lab wt = 3.80 kg

Snow sampling Date: 3-5-04

Site: C3C  
Initials: MD  
TH

Time: 12:00  
Air temp: 0°

Depth survey: BLUE

Wt. empty tube: 3.35

Depth of snow: ~~25~~ 19.5

Length of core: 19.5

Wt. tube + core: 3.96

Remarks: had a hard time getting the  
core because of rocks

Snow sample collection

Don't touch ORANGE

Hg samp remarks cedar branch  
in sample - pulled out. ice block in  
bottom, granular on top.

4.86 lab wt

Non-Hg samp remarks

Cap to cover ~~between C3A  
& C3C~~  
- Double bag - lab wt = 4.84 kg!



M edu M



Oak Ridge  
National Laboratory

Main Menu

- > Home
- > Consortia
- > Projects
- > Data Products

Projects Using

Below is a list  
of projects using  
this description of  
data. Click on the  
link to go to the  
project page.

ORNL Distributed  
Computing Center  
(DACC)

Overview

ORNL DACC

ORNL DACC

Metadata Editor

Core, EML, ISO-19115

SO data providers

http://maonline.net/ost/clarinet/home



# Informatics Approaches



- **Ontologies**

- **Ontology:** An explicit formal specifications of the terms in the domain and relations among them (Gruber 1993)
- Not simply a taxonomic structure, but explicit definition of concepts and relations
- **Examples**
  - ✦ GeneOntology- <http://www.geneontology.org/>
  - ✦ Marine Metadata Interoperability- <http://marinemetadata.org>

# How do we get at the *meaning* of data?



## Hg Ontology design

- Computer interpretable specification of concepts and relations among them
- Concepts: objects, processes, events

### Object:

Lake  
Catchment  
Bird  
Soil  
Sample

### Process:

Emission  
Deposition  
Methylation  
Bioaccumulation  
Biomagnification

### Event:

CAAA  
Storm  
Drought  
Sampling?

- Biomagnification is an *occurrence* and is *cumulative*
  - ✦ Occurs in BIOTA
  - ✦ Has a RATE
  - ✦ Is controlled by BIOACCUMULATION

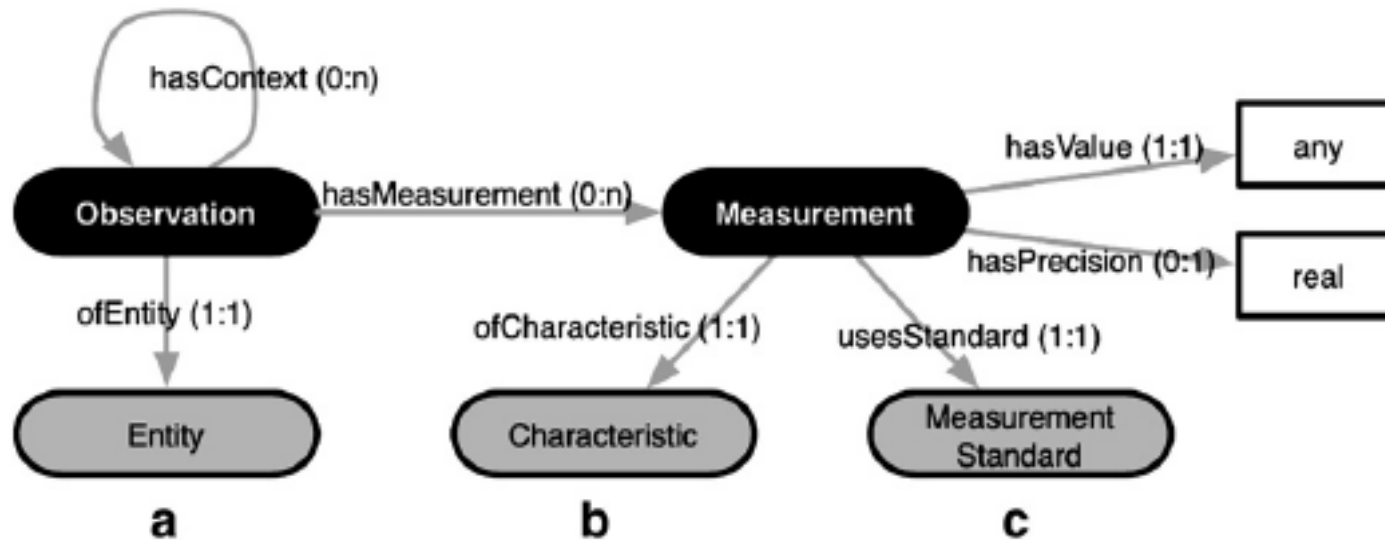
# OBOE- Extensible Observation Ontology



PHENOMENA

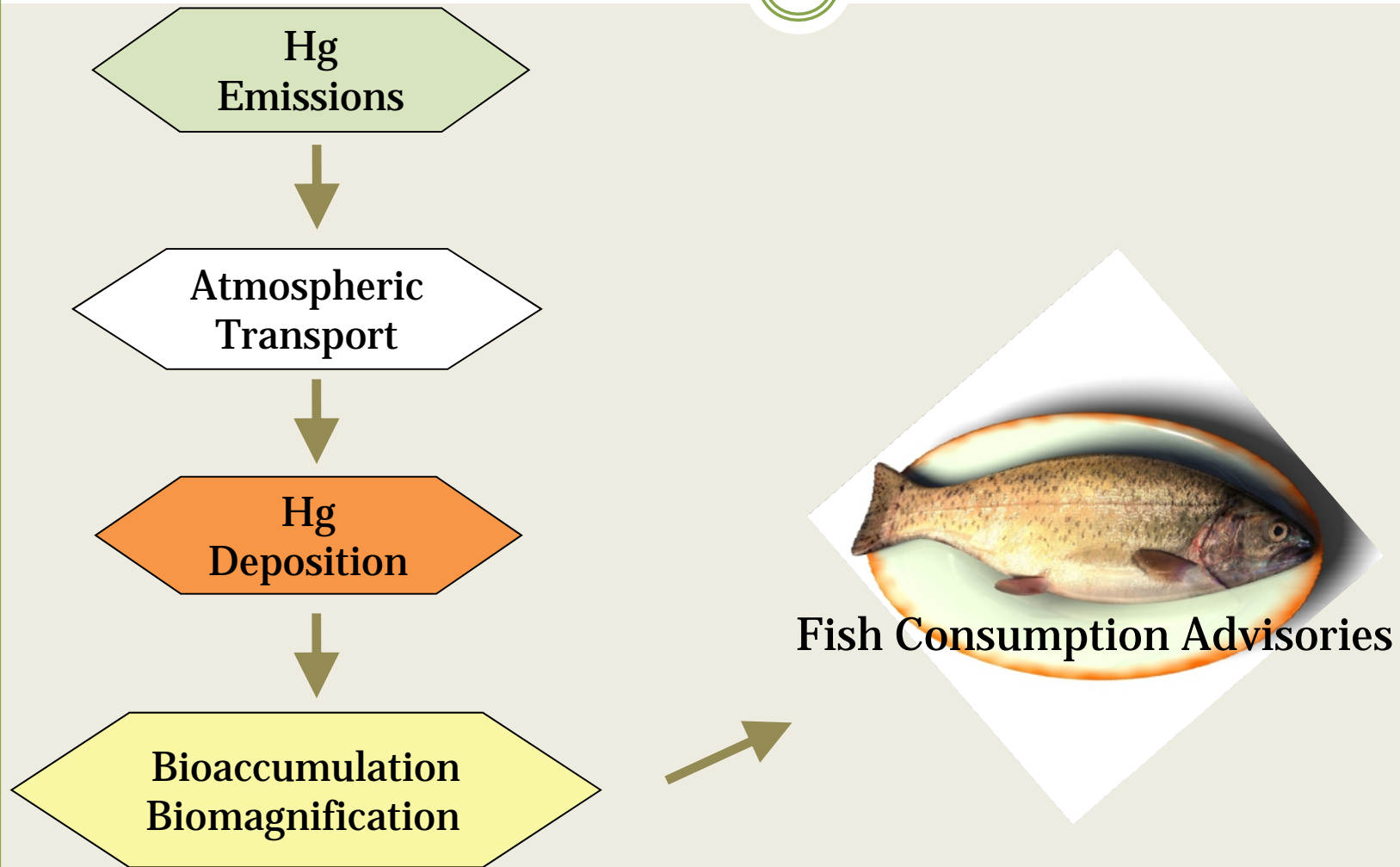
MEASUREMENT

INFERENCE

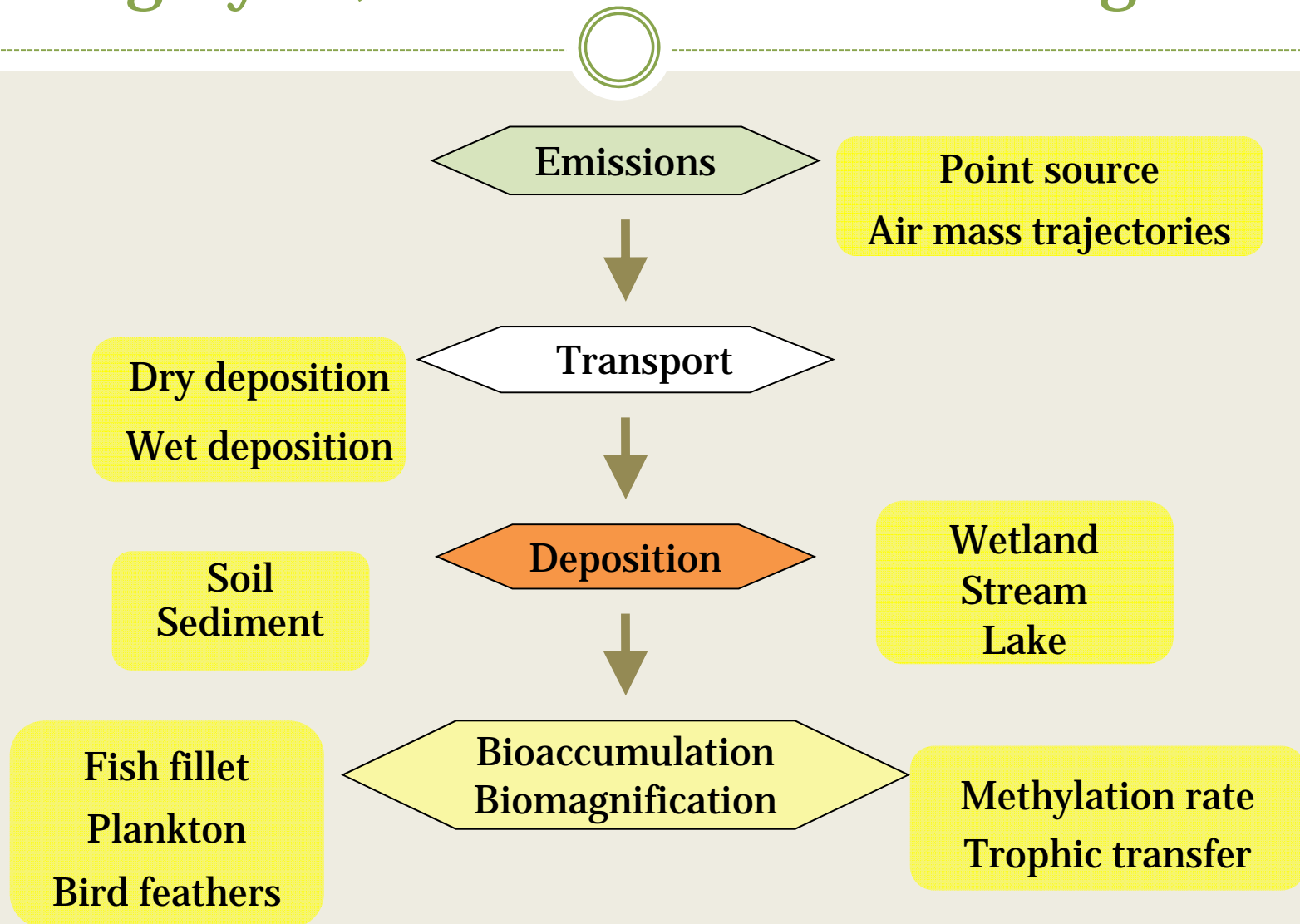


*From Madin et al 2007. An ontology for describing and synthesizing ecological observation data. Ecological Informatics 2(3):279-296*

# Hg Cycle, Briefly



# Hg Cycle, what are we measuring?

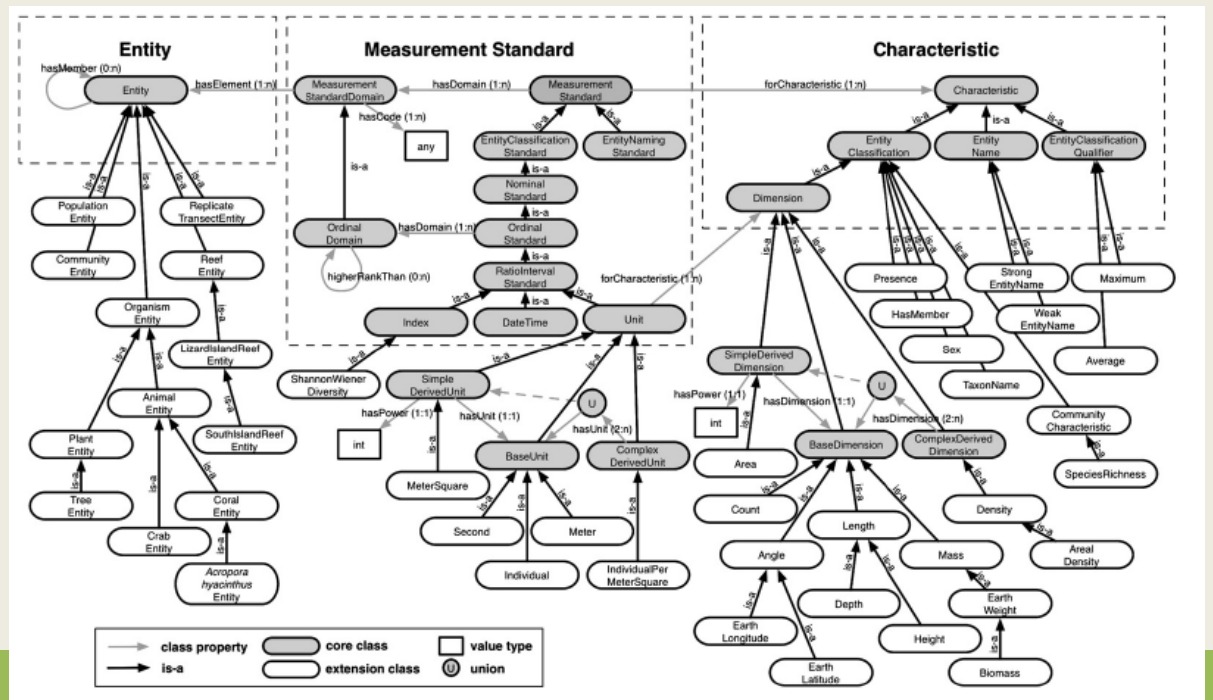


# Ontologies for Hg



## Under Construction:

- Using the Web Ontology language (OWL) and Protégé-  
protege.stanford.edu
- Will contextualize Hg data within the biogeochemical cycle and  
observation settings.



# Summary

- Do current metadata standards support comparison across datasets?

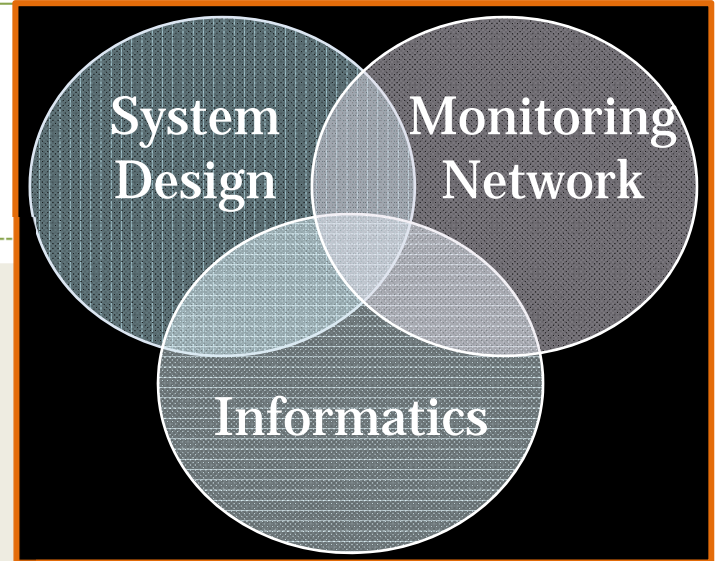
**Not yet, but the creation of an ontology can preserve metadata while creating relationships and linkages among disparate datasets.**

- What approaches should be adopted to enhance data repurposing?

**Metadata creation and management, public metadata archives, standardization of reporting.**

- Can the semantics of Hg research be standardized sufficiently to allow reasoning across large collections of disparate data?

**The tools are available, and have been successfully employed in other disciplines...**



# Acknowledgements



THIS work has been generously funded by the National Science Foundation IGERT Program under Grant 054494 and the Maine Water Resources Research Institute Grants Program.

