

# Human Centeredness

- Bill Sonntag
- Charlie Schweik
- Carol Hert
- **Eric Landis**
- Nancy Tosta
- Tyrone Wilson
- Steve Young
- Cliff Duke
- Mike Frame
- Doug Beard
- Sylvia Spengler
- Val Gregg

# What is covered/should be included in human-centeredness?

- The social issues that need to be included in ecosystem informatics decision making, including HCI, tech transfer, training, ...

# Collaboration

- What enables collaborative efforts?
  - Incentives
  - Rules
  - Self-realized values
  - Education
  - Others (time, training,
- What models (institutional design) exist in EI and elsewhere?
  - National bird count, Open source, meteorological, ....

# Disincentives to Collaboration

- They are there. What are they? Fear, privacy, cultural background,...
- Models of these are difficult to find as the projects probably failed at initiation.

# Broad question

- What needs to be in place to enable collaboration?
  - Many aspects
    - Training, education, user needs, standards, .....
- Tomorrow's assignment
  - Look at specific aspects and develop research questions.

# Modeling Breakout

## Modeling is important!

- Models are dynamic systems that need to change with probing, criticism, be supportive of deliberations and sensitive to policy.
- Models are hypotheses, and used to gain understanding of systems – they may not give a perfect or even “correct” answer.
- Models range from data intensive, complex to simple push button tools.

# Research issues:

- **Coupling diverse models:**  
different assumptions, definitions  
accounting of error (and propagations especially  
with introduction of multiple scales)  
Handle wide range of spatial and temporal  
scales
- **Visualizations** (results, also model structure,  
processes and influences)
- **Large data sets** and related performance  
challenges
- Creation of **software infrastructure** that  
supports writing transparent, flexible, reusable  
and credible models

# Considerations for building a modeling infrastructure

- Formal methods for evaluating the applicability of model uses, (including Bayesian, multi-attribute methods and game theory) weighing precision/realism/generalality.
- Comparison of models, including where they fail and their strengths along the lines of ensemble modeling in weather forecasting.
- Software engineering issues, including:  
Extensibility / flexibility, Open source/community software as a distribution method
- Sociology of models: use and collaboration

# Models should also include:

- applications and use of game theory and other decision making mathematical sciences to eco-informatics
- formal methods for evaluating the applicability of model uses, including techniques like Bayesian and game theory in the arena of eco-informatics, e.g. precision/realism/generalality
- those looking at multi-attribute decision-making

# Data Quality:

## Articulating research issues in eco-informatics decision- making

Larry Sugarbaker

Sherry Pittam

Kevin Gergely

Craig Palmer - presenter

Julia Jones - scribe

## Defining data quality

Data = error (noise) + signal (information)

Components of error are reproducibility and accuracy

Several sources:

- measurement error among human observers.
- instrument error and/or detection limits,
- natural variability

Question we addressed was not how to quantify error in primary data, but ...

**How can error estimates be incorporated into decision making?**

Decision-making typically based on combined datasets, from various sources.

Each data source has its own uncertainty, which are then combined in some unknown way when data sources or layers are combined.

**How would we communicate this uncertainty to decision makers?**

## Research questions for individual studies:

- How do errors arise in a study – can we list the steps at which errors might be produced?
- How should errors be measured at each stage? Are errors quantitative? Qualitative?
- How do errors occurring at various stages related to one another? Are errors compounded in the study? Or are they independent?
- How do we calculate errors in aggregated datasets (e.g. harvested ones)? For example, how do we validate the uncertainty estimates produced from integrating modeled values with field observations?
- How can uncertainty estimates be associated with particular alternative sets of actions that decision makers are evaluating?

## Research questions for sharing data on the web:

### **What does it mean to automate the management of metadata?**

#### Related questions

- Do downloadable data automatically include metadata on data quality?
- Can metadata be combined from multiple sources? How?
- Doesn't this generate a new measure of error for every data point, which is computationally challenging?
- Do standardized data formats help to simplify the problem of calculating the errors from combined datasets?
- Is uncertainty in some types of data more tractable than in others (e.g. standard format data like climate, hydrology [Clim-DB, Hydro-DB])?
- How can we identify forms of data that cannot or should not be used because of their effect on uncertainty?

# Information Integration

- Disclaimer
- Issues in information integration
- Technology for information integration
- Research Issues
- ... and another idea ...

# Issues in Information Integration

- Confidentiality
- Semantics
  - multiple definitions (including local terms)
  - multilingual
- Partnerships can help develop common vocabularies / semantics
- Citizen as client/user
- Local vs. national vs. international data
- Info exchange vs. integration

# Issues in Information Integration

- Description of information for proper use (including uncertainty)
- Integrating data of unknown or disparate uncertainty / science consistency checks / comparability
- Ethics of decision making --- how much / what to reveal
- How do you quantify semantic distance
- Creating semantic agreement beforehand is highly valuable

# Technology for Integration

- Web services
- Protocols for data collection (with QA) incl. definitions and measures
- Expert – analysis review (human in the loop / documentation)
- Wiki – to enable communities of practice (for semantic interoperability)
- Publication of best practices (standards/metadata, process/protocols)
- Indicators (core set process)
- Virtual data layers
- OASIS standards on web services
- W3C RDF usage (ontologies, rule sets)
- XML/RDF/OWL

# Research Questions

- Define / articulate dimensions of integration
- How do you quantify semantic distance?
- Integrating multiple ontologies?!
- How to promote modelling of documents (at doc creation)?
- How can we evaluate utility of varied data incl. qualitative and semi-quantitative data!
- Tools to support data integration!
- How can we elicit/evaluate tribal or other knowledge?

# Definition of Information Integration

- Mechanisms for reliable, transparent, authoritative data combination

# Ontology issues

- Uses: ontology as:
  - **metadata** over database(s)
    - Semantics for data, and cross-database integration
  - **standards** definition mechanisms
    - Cross-disciplinary connections
  - **terminology networks/taxonomies** (thesauri)
    - Search: pointers to data and associated info
    - Teaching/exploration of domain
  - support for (formal) **reasoning** systems
- Semantic environment:
  - Semantic Web: learning who else has Os that you can learn/steal from
  - The Grid:
    - Semantics: existing Os, termsets, etc.
    - Function : where can I find functionality? What must do do? What does it cost? When can I use it?
    - Data: where can I find data? What does it cost? Etc.

# Tools needed that require research

- **O building tools:**
  - **Pre-building** tools: tools to find related Os and termlists from the web, dictionaries, gazetteers, fact books, etc.
  - **Manual** building: interfaces for term entry by experts
  - **Automated:** tools to harvest info (from glossaries, domain text, existing metadata on 'orphan' databases, database tables, etc.)
  - **Mixed-initiative:** tools to merge existing Os
  - **Joint O building** support: negotiation support tools
- **O verification tools:**
  - **Internal:** tools that consider O structure, redundancy, etc.
  - **External:** tools that compare O to info:
    - relative to surrounding domain text, etc.
    - automatically finding and comparing to related Os on the web
- **O delivery tools:**
  - O services: what services needed? For whom? When?

# O-related phenomena to be handled

- **Incompleteness**: recognizing, recording, and warning of gaps in the O
- 'Vagueness': characterizing level of **granularity** of representation at each point/region in the O
- **Change and evolution**:
  - technology: versioning++; anything else?
  - representation 'theory': characterizing dimensions of change
  - future impact: manage and check for expected changes
- **Trustworthiness**: rating of source, whether human or not, as well as O acquisition tools/procedure
- **Controlled inconsistency** (microworlds): recognizing when animals can talk and handling exceptions