Eco-Informatics and Natural Resource Management

Judith Bayard Cushing
The Evergreen State College
Olympia WA 98505
USA
01-360-867-6652
judyc@evergreen.edu

Tyrone Wilson
National Biological Information Infrastructure
U.S. Geological Survey
Reston, VA, USA
01-703-648-4075
tyrone_wilson@usgs.gov

Alan Borning (University of Washington), Lois Delcambre (Portland State University), Geoff Bowker (Santa Clara University), Mike Frame (USGS/NBII), John Schnase (NASA), William Sonntag (EPA), Janos Fulop (Hungarian Academy of Sciences), Carol Hert (University of Washington-Tacoma), Eduard Hovy (University of Southern California), Julia Jones (Oregon State University), Eric Landis (Natural Resources Information Management), Charles Schweik (University of Massachusetts-Amherst), and Lawrence Brandt, Valerie Gregg, and Sylvia Spengler (National Science Foundation

ABSTRACT
Resource managers face significant information technology (IT) problems when integrating ecological or environmental information to make decisions. At this NSF/USGS workshop in December 2004 [1], university researchers, natural resource managers, and information managers articulated five IT problems areas facing ecology and environmental decision makers: policy making and implementation, data presentation, data gaps, tools, and indicators. To alleviate these problems, participants recommended informatics research in four IT areas: modeling and simulation, data quality, information integration and ontologies, and social and human aspects. Additional recommendations to assure cycles of innovation in the domain were addressed.

Categories and Subject Descriptors
H.2.8 [**Database Applications**]: Scientific Applications.

Keywords
Information integration, modeling and simulation, data quality, human-computer interaction, web services, eco-informatics.

1. INTRODUCTION
The 1998 PCAST report characterized bio-informatics as a biology and CS/IT cross-discipline, recognized the biodiversity-ecosystem nexus as an information enterprise, and envisioned analytical and synthetic capabilities among other foci in the next generation of NBII-2 information services [1]. Informatics tools for solving environmental challenges, e.g., global climate change, emerging diseases, decreasing biodiversity, and waning resources, are researched and developed under the *eco-informatics* rubric, but largely absent from university eco-informatics research is improved IT for natural resource management.

Participants in the December 2004 Eco-Informatics and Policy Workshop sought to address this gap. As domain problems and corresponding research areas were articulated, they found overarching issues. First, meeting natural resource management IT requirements is not simply a matter of adopting technology from related domains. Solving IT problems faced by natural resource decision makers requires, in addition to new research, innovation across the public and NGO sectors. This domain is similar to digital government IT research in that finding the right domain problem, distilling a range of research fruitful to all stakeholders, finding the right agency collaborators, and managing expectations are all critical. Researchers must also combine quantitative with qualitative information, and understand decision-making processes. If computer scientists and social scientists in the academy are not prepared for these challenges in addition to the demanding research outlined below, then eco-informatics in natural resource management will suffer.

2. THE PROBLEM SPACE
The problem space and required IT for eco-informatics policy making is broad because stakeholders include communities of interest as well as of place. Further, political aspects of the data necessitate equitable information access and mass customization, and distinguishing among measurements, indicators, and interpretations. Finally, metadata and validation are essential. More specific problems are categorized into five areas:

1) Policy problems include, but are not limited to, problems organizations encounter: 1) financing, production, and maintenance of tools and information; 2) use and possible abuse of tools or information, and their effectiveness; 3) cross-organizational sharing or not sharing, e.g., privacy, confidentiality policies, of IT tools or information; 4) communication, or lack thereof, of environmental management decisions grounded in eco-informatics-based analysis.

2) Data presentation problems arise from complex interactions between user needs and data – metadata, raw data, accuracy specifications, methods, documentation, and policy. Research required includes determining what information works best on which medium, cross-referencing and supporting data across presentations, representing time and change, using new media such as 3D or VR, and user task definitions. This area distills into two parts: presentation as mediator among users and data/metadata.

3) Geographic data gaps stems from a lack of needed data sets or access to them, disjunct data sets that require manipulation because of temporal or spatial gaps, an emphasis on adaptive management that out-paces data reliability, and a lack of IT professionals upon whom resource managers can call for expertise. Major issues include how to generalize fine-scale data containing gaps, and decision makers' and policy makers' sensitivity to uncertainty.

4) Tool problems involve 1) lack of a tool "clearing-house", 2) problems of data collection and new or different data types, 3) lack of user frameworks and development standards, 3) lack of tools to support metadata issues, and 4) social science issues of usage, sharing, and adoption.

5) Indicator problems exist because indicator definition, relevance, and value are neither well-defined nor well-communicated. Constituents may be uneasy with environmental measures, and data gaps adversely effect the reliability and trust that stakeholders place in indicators. The inherent ecosystem complexity further complicates this issue.

3. RESEARCH ISSUES

To articulate research issues, interdisciplinary approaches adopted by several current successful research projects were examined. Strategies for sustaining future research and case histories that exemplified the need for the research were identified as critical. Four research areas were identified:

1) Modeling and simulation research issues included: coupling diverse models, exploring model design values for diverse stakeholders, incorporating visualizations with model results, representing error and uncertainty, handling large data sets, and open source modeling infrastructure. An open-source, flexible, reusable modeling infrastructure, along with the social practices that sustain it, were seen as critical since it would allow researchers and decision makers to experiment freely with new models or change existing ones.

2) The problem of **data quality** is how to determine and communicate uncertainty to decision-makers for studies integrating multiple data sources. Methods are needed to mitigate error when creating and combining data sets, and to associate error with alternative decisions. NSF could publish metadata standards for all grants, rather than just certain programs, and make metadata obligatory for all data sets. Whether uncertainty associated with data synthesis influences resource management, whether decision-maker perceptions of the value of findings from synthesized data, and how synthetic studies compare in courts of law to other "expert testimony" remain open questions.

3) Information integration involves mechanisms for reliable, transparent and authoritative data combination. Associated research issues include: defining the dimensions of integration; quantifying semantic distance; integrating multiple ontologies; promoting document modeling; evaluating utility of qualitative and quantitative data; tools to support data integration; and evaluating knowledge from non-traditional sources. **Ontologies** are useful in providing semantics over databases, making cross-disciplinary connections, and producing thesauri. Tools to build, verify and deliver ontologies require considerable research. Other research areas are understanding gaps and inconsistencies in ontologies, trust and verification of ontology content, and ways to handle change in material represented by onotologies that go beyond versioning.

4) Social and human aspects of eco-informatics and policy include: collaboration in IT tool development and information sharing among decision makers, measuring success and determining appropriate institutional designs and incentives or disincentives; human-computer interaction; management practices, graduate education and data management training. Advancing the eco-informatics agenda hinges on both new technologies, and new understandings of how information infrastructures inter-relate between individuals, organizations, communities, disciplines, information resources, and tools.

4. RECOMMENDATIONS

Considerable attention should be paid to assuring a cycle of innovation from research to prototype, to development and commercialization, and finally to deployment and evaluation (and back to research). Differing, non-overlapping missions and reward systems built into different agencies make it too easy to lose momentum at any of these stages. Longer funding cycles are needed to elicit requirements and integrate these into a research agenda, and then enter into an "agile" software cycle of develop, evaluate, and deploy. Funding agencies must work together and with principal investigators, information managers and decision makers in sustaining and encouraging innovation, research and development, and education and training. How would researchers find collaborators so they can best understand resource problems, extract the research issues, and test their prototypes? How might research results and prototypes make their way to resource managers as IT deployed in field offices? How would new products be evaluated, and an understanding of strengths and weaknesses, inform new research?

Considerable attention must be paid to constant reprioritization of the research agenda, since the sheer number, breadth and complexity of problems and potential solutions suggested herein dictate decades of research – while species and ecosystems disappear at an increasing rate. Thus, assuring the development of tools that promise, through extensibility, application to a wide range of problems is critical, as is keeping a range of research projects in the pipeline – from highly theoretical and generalized, to working prototypes developed by researchers and resource managers, to deployment experiments.

5. ACKNOWLEDGMENTS

The authors thank presenters[1] and participants in the Eco-Informatics and Policy Workshop. For their helpful comments we thank Aaron Ellison of Harvard University and Elaine Hoagland of The National Council for Science and the Environment. Finally, we acknowledge funding from the National Science Foundation (NSF IIS 0505790).

6. REFERENCES

[1] For the full workshop website, including all presentations, the final report, and a listing of all participants, see http://www.evergreen.edu/bdei.

[2] http://clinton3.nara.gov/WH/EOP/OSTP/Environment/html/teamingcover.htm

[1] Presenters who are not also listed as report authors include: C. Marie Denn (USPS), Richard Guldin (USFS), Stephen Jensen (EU), Paul Klarin (Oregon Coastal Management), Ron Li (Ohio State University), Molly O'Neil (Harry Reid Center for Environmental Stdies), Phil Rossignol (Oregon State University), Mark Simonson (Puget Sound Regional Council), Larry Sugarbaker (NatureServe), Nancy Tosta (Ross&Associates), Dawn Wright (Oregon State University).