Project Description

BDEI PI's Workshop:

Reporting Results & Research Prospects of Planning and Incubation Grants

Introduction

This proposal seeks funds for a workshop (currently scheduled for February 11, 2003) to bring together Principal Investigators of NSF's nascent Biodiversity and Ecosystem Informatics (BDEI) Program **to report research results of BDEI planning and incubation grants to program directors in NSF (BIO, CISE and GEO), NASA and USGS.** We aim to use those research results as a basis for articulating **future research prospects, in particular fundamental CS/IT research directions needed to solve major BDEI research questions.** We will also aim to **identify areas where focused applied CS/IT research would yield considerable immediate benefit in BDEI domain research**. Secondary aims of the workshop are:

(1) to identify benefits and value to the computer science research community of BDEI-focused CS/IT research,
(2) to articulate impediments and incentives for computer scientist involvement in BDEI and develop a plan to pique the interest of the CS/IT research community,
(3) to carry out the initial steps of the CS/IT outreach plan,
(4) to identify specific research areas where interdisciplinary and disciplinary educational outreach in CS/IT and BDEI, or cross-disciplinary training of early-career investigators, have particular long term impact.

Because BDEI is potentially a domain where obvious societal impact of research could be used to attract women and other under-represented groups to CS/IT research careers[1], we will make a particular point of considering potential impact of past and future research on CS/IT education. The workshop report will also address how early-career investigators can be attracted to this domain and how training can be most effectively targeted for them.

We will publish and present to computer scientists the results of the current work and CS/IT benefits of the research. A special issue of the *Journal of Intelligent Information Systems* (JIIS) that will publish in one volume papers by BDEI Investigators has already been proposed, and the PI's will use the workshop to fine-tune that publication and find an appropriate outlet for showcasing work not appropriate to or as yet ready for the JIIS issue.

Invitation to the workshop will be by invitation only; we have confirmed the date with all BDEI PI's and with agency program directors in NSF (Larry Brandt and Valerie Gregg, EIA and DG; Sylvia Spengler, BIO; Maria Zemankova, CISE), NASA (John Schnase, Jim Smith, and Woody Turner) , and USGS (Mike Frame and Gladys Cotter). Though we will keep the workshop small (less than 35), focused, and interactive, included in that group will be a few attendees of the original workshop as well as respected experts in computer science and the domain. These persons will add balance and perspective to the discussion of research results and directions. Susan Stafford, Dean of the University of Minnesota College of Natural Resources and Chair of the Long Term Ecological Research Center Information Managers Executive Committee, has agreed to attend. The workshop organizing committee is currently identifying other participants.

[1] Margolis and Fisher, *Unlocking the Clubhouse: Women in Computing,* MIT Press, 2002.

The workshop will be held at NASA Headquarters in downtown Washington, D. C., at no cost to NSF; space has already been reserved. In this proposal, we seek travel and lodging expenses for participants whose agency, institution or company can't cover those costs, as well as meals and breaks.

Principal Investigator	Title
K. Beard-Tisdale U Maine	Event And Process Tagging For Information Integration for the International Gulf of Maine Watershed
K. Bergen U Michigan	Radar Remote Sensing of Habitat Structure for Biodiversity Informatics
G. Bowker UCSD	Designing an Infrastructure for Heterogeneity of Ecosystem Data, Collaborators and Organizations
J. Clark Duke	Computation and Uncertainty in Ecological Forecasting
J. Cushing Evergreen State College	Spatial Data Infrastructure for Ecological Research (Planning Grant)
A. Dickerman VPI	Bioinformatic Prediction of Functions of Unculturable Microbes in Ecosystems
P. Doruska U of Arkansas	Quantifying Forest Ground Flora Biomass and Diversity Using Close-Range Remote Sensing
P. Flikkema NorthernArizonaU	Reconfigurable Wireless Sensor Networks for Dense Spatio-Temporal Environmental Monitoring
S. Gauch U of Kansas	Biodiversity Information Organization using Taxonomy [BIOT]
G. Henebry U of Nebraska-Lincoln	Spatio-temporal Models of Biogeophysical Fields for Ecological Forecasting: A Cross-Disciplinary Incubation Activity
M. Lane Acad Nat Sci of Phila	Overcoming nomenclatural complications while searching in a distributed database environment: One step toward true interoperability
M. Musavi U of Maine	Planning Workshop on Biodiversity and Ecosystem Informatics for the Indian River Lagoon, Florida
R. Stevenson U of Mass Boston	The Eco Flight Simulator: Visualizing landscape patterns, ecosystem processes and biodiversity information
F. Stevenson U of Mass Boston	Planning a Community Science Approach to Biodiversity Monitoring: Extending the Spatial and Temporal Scales
F. Villa U of Maryland	Towards an Operational Semantics of Biological Diversity: Integrating Structure and Function in a Web-accessible Knowledge Base

Table 1: NSF BDEI Research Initiation and Planning Projects Awarded in 2001

Background

In June, 2000, an NSF-NASA-USGS sponsored workshop on biodiversity and ecosystem informatics brought biologists, ecologists and resource managers together with computer scientists at NASA Goddard to identify the CS/IT research issues that impede biodiversity and

ecosystem research or the development of a national ecosystem network. Their report[2] sparked the National Science Foundation to issue a cross cutting Dear Colleagues call for proposals (due July 2001) for research planning and initiation[3]. Fifteen awards were made that fall. The principal investigators of these grants came together in the spring of 2002 for a first meeting at the Digital Government annual meeting. Thus, a major goal of the June 2000 Workshop was attained – an initial generation of interest in BDEI research from the CS/IT and Ecology Research Community. Table 1 shows the PI's and titles of NSF-funded BDEI projects, and Table 2 shows a categorization of those projects according to topic area and PI discipline.

Projects \ Categorization	Data Integration Ontologies Data Modeling Metadata	Spatio-Temporal / Remote Sensing	Models, Monitoring & Forecasting (PtoP)	CHI & Collab. Work	Work-flow	PI Discipline	
						ECO	CS
Community Science (Stevenson)	X	X	X	X	X	X	C
Ecosystem Workshop (Musavi)	X					X	X
Taxonomy (Gauch)	X					X	
Nomenclature (Lane)	X					X	
Operational Semantics (Villa)	X					X	
Event & Process tagging (Beard)	X				X	X	L
Spatial Infrastructure (Cushing)	X	X		x		X	
Vis. Patterns & Processes (Stevenson)		x	X			X	X
Spatio-Temporal Models (Hennebry)		X	X			X	X
Reconfig.Wireless Sensors (Flikkema)		X					ee
Radar Remote Sensing (Bergen)		X	X			X	ee
Computation & Uncertainty (Clark)	x		X			X	X
Microbial Fcn Prediction (Dickerman)	X		X			X	
Data,Collab,Organizations (Bowker)				X			
Biomass&CloseRangeSens'g (Doruska)		X	x			X	
SEEK (ITR)	X					X	X

Table 2. A Preliminary Categorization of BDEI Projects according to CS/IT Research Area and PI/coPI discipline. Discipline codes are as follows: ECO= Ecology; CS=Computer Science. Within CS, related disciplines are noted as follows: C = computer supported cooperative work; L = Library Science; ee=electrical engineering.

[2] See *Research Directions in Biodiversity and Ecosystem Informatics*, http://bio.gsfc.nasa.gov, report of the NSF, USGS, NASA Workshop on Biodiversity and Ecosystem Informatics, June 22-23, 2000.
[3] See NSF Dear Colleagues Letter, available at http://bio.gsfc.nasa.gov/bddcl.pdf

The BDEI PI's met at the Digital Government PI's Meeting in May 2002, and each presented a short summary of his or her research. Short papers and slides are available on the DGO website[4]. In addition, Cushing convened a panel at the August VLDB 2002 Conference (K. Bergen, J. Kennedy, R. Miller and Y. Ionnidis) where panelists were asked to assess whether the NSF-funded BDEI research constituted CS/IT Research, Technology Transfer, or Application Development. Panelists' slides and a summary of the ensuing discussion are available on the VLDB web site[5].

Organizing Committee

We have formed a small organizing committee for the workshop, led by Judy Cushing, the PI of this proposal.

- Judy Cushing, The Evergreen State College and BDEI PI
- John Schnase, NASA Goddard Space Flight Center
- Woody Turner, NASA Headquarters
- Eric Landis, Editor of the June 2000 BDEI Workshop Report
- Kathleen Bergen, University of Michigan and BDEI PI

We have also been in communication by email and phone to determine the scope and structure of the workshop, and seeking advice from the following NSF and USGS program officers:

- Larry Brandt, Program Director, Digital Government (EIA)
- Sylvia Spengler, Program Director, Biological Databases and Informatics, Research Resources (BIO)
- Maria Zemankova, Program Director Information and Data Management (CISE)
- Mike Frame, Deputy Center Director for the Center for Biological Informatics, USGS

The organizing committee aims to appoint three or four facilitators from among the 14 PI's for each of 3 or 4 subgroups of projects, based on the preliminary categorization of research. The main categories that we expect will comprise panels are: semantic data integration (incl. ontologies, data models, and metadata); spatio-temporal and remote sensing data; and models, monitoring and forecasting. A few projects involve community organization issues and user interfaces, and scientific research process workflow. Results of these projects will likely be presented in a panel whose topic deals with whether change in ecosystem and CS/IT research and development process is advisable, and if so how to effect that change.

The research category facilitator (panel moderator) will be responsible for defining the area and identifying area research questions, conducting a panel session during the workshop of those researchers, facilitating or arranging facilitation of the corresponding breakout group, and completing a preliminary report of research directions for that area.

[4] See http://www.dgrc.org/dgrc/dgo2002 and http://www.dgrc.org/dgrc/dgo2002/html/program-apr18-details.htm#theme1.
[5] See http://www.cs.ust.hk/vldb2002 and http://www.cs.ust.hk/vldb2002/program-info/panels.html.

Workshop Structure

We plan on an initial keynote or capstone presentation at the start of the workshop to set the stage and ground the subsequent research presentations from the CS/IT point of view. The major vehicle for content presentation, however, will be research presentations by the BDEI PI's. PI's will be asked to give equal weight to four things: research results, potential future work, relevance of that future work to both BDEI and CS/IT, and opportunities inherent in that future work for educational outreach in CS/IT and BDEI.

Prior to the workshop, participants will be provided with short summaries of each of the research projects; at breakfast and breaks, we may organize informal poster sessions of the BDEI research. We have decided that panel format for research presentations, with a moderator who organizes the panel prior to the workshop and strictly enforces time limits is the best way to present research results. Our aim is to focus each panel by first defining and explaining the CS/IT research area, then taking each grant activity as an example of viable research in that area, and finally summarizing next research steps.

Following the morning panel presentations, breakout sessions will work with panelists to define research agenda for each of the three major focus areas. Panel facilitators will facilitate these discussions. After reports by the working group facilitators, a panel of expert but independent scientists will comment on the results reported. A final discussion will focus on how to assure funding for future BDEI CS/IT research (whether to recommend a new program or provide direction to existing programs), ways to involve computer scientists and how to foster effective cross disciplinary collaboration. Following the workshop, panel facilitators and workshop organizers will briefly meet to assure timely deposition of written sub-committee reports.

On the day prior to the workshop, the organizers will meet to fine tune the schedule and panel sessions (4-5:30). A dinner meeting of PI's and panel moderators will follow (6:30-9), where each panel reviews its objectives and results, and walks through their presentations. A tentative schedule for the workshop follows:

Monday, February 10

~4-5:30	Organizing Committee plus Panel Moderators meet
~6:30-9	BDEI PI's meet for dinner at the hotel (Downtown D.C. Hilton) dry run of panel presentations; preliminary brainstorm of research result summary & agenda

Tuesday, February 11

8:00	Breakfast; Posters (possibly) Available
8:30-9:00	Welcome and Keynote
9:00-10:30	Panel: Semantic Data Integration (Musavi, Gauch, Lane, Villa, Beard)
10:30-10:45	Break
10:45-12:15	Panel: Spatio-Temporal Data / Remote Sensing (Cushing, Stevenson,Hennebry, Flikkema, Doruska)
12:15-1:15	Lunch
1:15-2:15	Panel: Modeling, Forecasting (Bergen, Clark, Dickerman)
2:15-3:00	Summary Panel: Putting it into Practice: Research, Community, Systems, Workflow (Stevenson, Bowker)
3:00-3:15	Break

3:15-4:00	Breakout Sessions – Confirmation of Research Agenda Laid out by Panel.
4:00-5:00	Expert Panel Report and Summary Discussion Particulars education and training to be included in discussion.
5:15	Final Remarks and Farewells
5:30-5:45	Panel Moderators, Expert Panel and Workshop Organizers exchange notes

Workshop Results and Conclusions

We see three main results from the workshop:

1. A more detailed CS/IT Research Agenda and Funding Recommendation than was possible from the June 2000 Workshop Report, before planning and initiation grants were awarded. This agenda will include descriptions of previous research, and pointers to additional information about that research. Specific recommendations about funding sources for this work will also be included. This report will be first published on the workshop web site and then later archived at NASA along with the previous BDEI workshop report.
2. Joint publication of BDEI research efforts, including the special issue of JIIS.
3. Outreach to the CS/IT community on benefits of BDEI research and advice on how to proceed, plus a plan of how to continue the dialog about BDEI (and other scientific related research) that was begun at VLDB 2000.

Impact of the Proposed Workshop

We expect the research resulting from workshop to impact three areas: Biodiversity and Ecosystem Domain Research, CS/IT Research, and CS/IT and ecology education.

Biodiversity and Ecosystem Domain Research. The June 2000 BDEI Research Directions Workshop Report showed unequivocally that BDEI research is needed to supply CS/IT infrastructure for critical biodiversity and ecosystem research. The BDEI research would have considerable albeit secondary societal impact when biodiveristy and ecosystem research is undertaken as has suggested by presidential committees (PCAST and PITAC[6]), the Office of Science and Technology Policy's Committee (OSTP/CENR[7]) and the National Research Council[8]. The NSF Dear Colleague letter states this potential impact clearly:

> In order to effectively address the highly complex environmental science questions of the 21st century, scientists, resource managers, decision makers, and interested citizens will need access to an entirely new generation of computer systems, tools, and capabilities. We currently lack the capability to effectively deal with the challenges of collecting, maintaining, analyzing, and understanding the huge amounts of biodiversity and ecosystems data that have been collected over the past 200 years - not to mention the new data which are being collected each day.

---

[6] See reports of the President's Committee of Advisors on Science and Technology (PCAST) Panel on Biodiversity and Ecosystems http://www.ostp.gov/Environment/html/teamingcover.html and the President's Information Technology Advisory Committee (PITAC) http://www.itrd.gov/ac/report.

[7] See report of the Office of Science and Technology Policy's Committee on Environment and Natural Resources (OSTP/CENR) http://www.ostp.gov/html/0076.html .

[8] See the National Research Council's Committee report on Grand Challenges in Environmental Sciences http://books.nap.edu/books/0309072549/html/R1.html#pagetop.

> These challenges reflect the fundamentally complex nature of biodiversity and ecosystems data (molecular and genetic diversity, species diversity, population and community-level diversity, ecosystem diversity); extremely large and widely distributed data sets in disparate forms and formats (including irreplaceable data that are not currently digital); and the need to support analysis and interpretation at varying spatial and temporal scales, including *in-situ* (wireless) computation and information access.
>
> It is expected that this call for proposals will push the boundaries in information representation, acquisition, integration, analysis, synthesis, access, use and long-term preservation. In many cases, wholly new approaches to geospatial and temporal data management will be required, as well as advances in areas such as computer-mediated collaboration, modeling and visualization, knowledge discovery, data mining, and remote sensing.
>
> Meeting these challenges will benefit a wide range of information domains and directly affect the future health and sustainability of biological diversity and our natural environment, including such crucial issues as species and habitat loss, invasive species, emerging diseases, effects of global change on ecosystem processes, and the restoration of lost or damaged ecosystems. It could also ultimately provide us with the advanced capabilities needed to support true "ecological forecasting," thus proving of immense value to the successful execution of government agency missions.

If the workshop report and outreach successfully enables and encourages CS/IT researchers to involve themselves in BDEI research, and if funding opportunities required to carry out that research are subsequently made available, this impact will occur.

CS/IT Research. The impact of BDEI research on CS/IT Research is as yet less clear, but has been suggested by the panel presentations at the Very Large Database Conference in August 2002. CS/IT researchers (Miller and Ioannidis) not currently involved in BDEI research stated that BDEI research problems are consonant with the current CS/IT research agenda, and that the state of the CS/IT art could be advanced were CS/IT researchers to undertake BDEI research. The reasons for this include: considerable complex data in the public domain (unlike typical industry data) and a knowledgeable ecological information management user community with whom CS/IT researchers could collaborate. The BDEI PI's workshop aims to identify if and how BDEI research would forward CS/IT research agenda. While there are perceived barriers to conducting this research, the workshop report and outreach will address those perceived barriers and determine the extent to which they are real or perceived, and work to dispel barriers and misperceptions.

CS/IT and Ecology Education and Training. In *Unlocking the Clubhouse* Margolis and Fisher pay considerable attention to the fact that women are attracted to and sustained in CS/IT education if they understand the positive societal impact of the work they will do. Biodiviersity and Ecosystem Informatics has clear, strong societal impact, as is articulated in There is some evidence that other underrepresented groups will similarly be affected. The National Research Council's Committee on Grand Challenges in Environmental Sciences has recently identified the need for fundamentally increased understandin g of biodiversity and ecosystem functioning as one of the eight "Grand Challenges in Environmental Science" facing our nation and the world. Surely, this provides considerable societal impact and can be made to provide exemplars of CS/IT applications that matter in significant ways.

Several current BDEI grants address how ecology researchers, computer scientists and information technology researchers can be educated and trained to best work together, and also how new CS/IT can enable lay public involvement in biodiversity and ecosystem research. Thus, results reported by the workshop will consider these critical issues.