

2015 NSF Workshop on “Intelligent Systems Research to Support Geosciences and the EarthCube Mission”

Workshop Objectives

This workshop will identify innovative cyberinfrastructure that pushes the envelope on intelligent and information systems research that will enable advances in our understanding of the Earth system. Participants will discuss how to tackle problems in heterogeneous data integration and visualization (e.g., hand-made sketches, aerial imagery, field-data repositories, stakeholder interviews), ontological reasoning with scientific metadata and mathematical models (e.g., representing uncertainty, simulation predictions, evolving theories). Additionally, participants will be asked to consider potential uses for intelligent assistants that make scientists more efficient and that facilitate information sharing, collaborative workflow design and management to support data analytics, and sophisticated machine learning techniques to analyze geosciences data. Addressing these challenges requires getting together researchers from geosciences and information systems. The workshop will catalyze a community and research agenda in the emerging area of Discovery Informatics grounded on geoscience requirements.

The recent NSF EarthCube Initiative (<http://www.earthcube.org>) opens an opportunity for collaborative research on novel information systems enhancing and supporting geoscience research efforts. EarthCube’s goal is to enable geoscientists to address the challenges of understanding and predicting a complex and evolving Earth system by fostering a community-governed effort to develop a common cyberinfrastructure to collect, access, analyze, share and visualize all forms of data and resources, using advanced technological and computational capabilities.

The goal of the workshop is to synthesize a vision and needs for intelligent systems research that will provide new capabilities to advance geosciences. In geosciences, the workshop will cover requirements from Earth, ocean, polar, and atmospheric and spatial sciences. In intelligent systems, the workshop will include participation from fields such as information integration, machine learning, knowledge representation, social computing, visualization, and intelligent user interfaces. The workshop will be informed by existing cyberinfrastructure efforts that support the geoscience community. The workshop will serve as a bridge to get these communities talking to one another and explore collaborative research.

The workshop will provide a forum for geosciences and computer scientists to formulate a research agenda for intelligent systems that will result in fundamental new capabilities for understanding the Earth system.

Workshop Co-Chairs

Dr. Yolanda Gil is Director of Knowledge Technologies and Associate Division Director at the Information Sciences Institute of the University of Southern California, and Research Professor in the Computer Science Department. She received her M.S. and Ph. D. degrees in Computer Science from Carnegie Mellon University. Dr. Gil leads a group that conducts research on various aspects of Interactive Knowledge Capture, including intelligent user interfaces, knowledge-rich problem solving, and the semantic web. An area of recent interest is collaborative large-scale data analysis through semantic workflows. She initiated and chaired the World Wide Web Consortium (W3C) Provenance Group that led to a community standard in this area, and is now engaged with the Open Geospatial Consortium (OGC) to investigate its use for geospatial data. Dr. Gil has served in the Advisory Committee of the Computer Science and Engineering Directorate of the National Science Foundation. She is Chair of ACM SIGAI, the Association for Computing Machinery's Special Interest Group on Artificial Intelligence. She was elected Fellow of the American Association of Artificial Intelligence (AAAI) in 2012. Dr. Gil has been an active participant of EarthCube since its inception. She was lead PI of the Workflows Roadmap activity, and was co-

chair of the End User Workshop for Early Career Researchers. She is PI of the GeoSoft Building Block award for fostering sharing and reuse of geoscience software, is in the Steering Committee of the EC4 Research Coordination Network for field data science, is senior personnel of the iSamples Research Coordination Network for organizing and digitizing physical sample collections. She serves in the newly instituted EarthCube demonstration governance as Chair of the Technology and Architecture Committee and as its elected representative to the EarthCube Leadership Council. More information about research activities, scientific publications, and a full CV are available at <http://www.isi.edu/~gil/>.

Dr. Suzanne A. Pierce is an Assistant Professor of Research with the Environmental Science Institute in the Jackson School of Geosciences. She is also the Assistant Director of the Digital Media Collaboratory in the Center for Agile Technology at The University of Texas at Austin. A trained hydrogeologist with a focus on participatory deliberation, Dr. Pierce has prior professional background as a Scientist with Sandia National Laboratories and as the Environmental Manager for one of the world's largest metals mines. Dr. Pierce adopts a scholar-practitioner approach to integrate science-based information with human organizational systems for application to groundwater management and energy-water problems. Resultant decision support systems link participatory modeling with simulation, optimization, and multi-stakeholder concerns. Current projects include development of hydroinformatics for sustainable aquifer yield in Central Texas and South Australia, along with creation of the ENCOMPASS cyberinfrastructure for a geothermal basin in the Atacama Desert of Chile. ENCOMPASS is a scientific platform for sharing data, algorithms, and educational modules that has been funded by the Fulbright Nexus program and the Longhorn Innovation Fund for Technology. Dr. Pierce was an invited participant in the EarthCube End User Workshop for Early Career Researchers, and was selected to be a member of the Early Career Advisory Committee of the EarthCube GeoSoft project for software stewardship in geosciences. More information is available at http://www.jsg.utexas.edu/researcher/suzanne_pierce/.

Workshop Agenda

Participants will be asked to contribute materials prior to the workshop, such as current interests, position papers, and relevant publications. These materials will be used for designing the workshop schedule as well as to prepare the workshop report.

In order to ensure meaningful interactions of the geoscientists with the computer scientists, there will be two sessions organized as a “World Café” (http://en.wikipedia.org/wiki/World_Caf%C3%A9_%28conversational_process%29). The basic idea of a “World Café” is to have a set of round tables each with a theme assigned to it and a participant assigned to that table, and the rest of the participants rotating around the tables at fixed time intervals. So in the first “World Café” session the tables will have geoscience themes and the geoscientists will be assigned to the tables, and during the session the computer scientists will rotate around the tables. In the second “World Café” session, it will be the computer science themes that will be assigned to the tables and the geoscientists will rotate around. This will create appropriate opportunities for interaction across the disciplines. These sessions will be followed by plenary discussions on the themes that arose from each of the round tables.

The workshop will end with a presentation open to NSF Program Directors that will summarize the outcomes of the meeting, and will serve as the outline for the workshop report.

Workshop Report

Salient themes arising from discussions at the meeting will be articulated in detail in the final workshop report. An outline of the report will be discussed at the workshop, with writing assignments. A first draft of the report will be drafted by within a month of the workshop. Participants will be asked to contribute to sections corresponding to discussion topics they led or contributed to. The report will then be edited and finalized by the Co-Chairs.

An important result of the workshop will be a journal article co-authored by workshop participants and based on the workshop report. The topic of the article will be the challenges and opportunities for intelligent systems in geosciences.

Prior Related Workshops

Several workshops have been organized in recent years on topics relevant to the proposed workshop, although the topic of intelligent systems for geosciences is itself new.

Several NSF workshops have been held as part of the EarthCube program (<http://www.earthcube.org/page/end-user-workshops>). These workshops, called End User Workshops, focused on the vision, challenges, and requirements of specific communities within geosciences (e.g., geochronology, ocean 'omics, critical zone, sedimentology, stratigraphy, etc). They were designed to develop a vision and associated requirements from a range of disciplines in the geosciences. None of the workshops focused on developing a vision for advancing cyberinfrastructure, which is the focus of the proposed workshop which is focused in particular on intelligent systems for geosciences. The reports of the EarthCube End User workshops will inform the proposed workshop, and selected participants will be invited to join the proposed workshop.

An NSF Workshop on Discovery Informatics, held in February 2012 in Arlington VA (<http://www.discoveryinformaticsinitiative.org/diw2012>). Dr. Gil was Co-Chair of the workshop and helped organize several symposia for the Artificial Intelligence community. That workshop did not have a specific focus on geosciences, but will inform the proposed workshop activities. A final report from that workshop was published in the Web site, as well as a slide presentation that was given at NSF and in several scientific meetings. A recent article in *Science* describes the results from that workshop [Gil et al 2014].

Several other workshops have been held over the past few years that can A series of workshops on the topic of Cyber-enabled Discovery and Innovation include the NSF Symposium on Cyber-Enabled Discovery and Innovation, held at Rensselaer Polytechnic Institute on Sept. 5-6, 2007 (<http://www.rpi.edu/nsfcdi/program.html>), the NSF workshop on data mining and cyber-enabled discovery for innovation, held at the University of Maryland on Oct. 10-12, 2007 (<http://ebiquity.umbc.edu/blogger/2007/08/05/nsf-workshop-on-datamining-and-cyber-enabled-discovery-for-innovation/>), and the SCo7 session on supercomputing and CDI, held on Nov. 13, 2007 (http://sco7.supercomputing.org/schedule/event_detail.php?evid=11289). Many workshops have been held concerning scientific challenges for cyberinfrastructure, including the NSF Workshop on Cyberinfrastructure for the Atmospheric Sciences in the 21st Century, held on June 2004 (http://netstats.ucar.edu/cyrdas/report/cyrdas_report_final.pdf), and the NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences, held at the San Diego Supercomputing Center on March 15-17, 2005 (<http://www.sdsc.edu/sbe/>). Other workshops include the NSF Workshop on Knowledge Management and Visualization Tools in Support of Discovery, held in Indiana University on March 2008 (<http://vw.slis.indiana.edu/cdi2008/workshop1.html>), the NSF Innovation and Discovery Workshop: The Scientific Basis of Individual and Team Innovation and Discovery, held in August 2006 (www.nsf.gov/pubs/2007/nsf0725/nsf0725.pdf).

A relevant workshop series that has been supported by NSF funds is the Climate Informatics Workshop series (<https://www2.image.ucar.edu/event/ci2014/>). The focus of the workshop series is the presentation of research results submitted by the community, rather than the development of a research roadmap.

These and other relevant workshops will be discussed at the meeting and summarized in the final workshop report.

Workshop Location

The workshop will be held March 26-27, 2015. The workshop location will be the Westin Arlington in Arlington, VA, a location near NSF that will facilitate participation of NSF Program Directors, as well as attendance from other government agencies.

Travel and Reimbursement

Travel expenses will be reimbursed when receipts are provided and expense is within reason.

The following will be reimbursed with submission of original “itemized” receipts (if the purchase is done online, a soft copy will suffice):

1. Airfare – Flights must be booked as a round trip from the home city to Washington, DC, be reasonably priced, and have economy class fare. You **MUST** use U.S.A. flight carriers as part of the Fly America Act. If your travel plans involve other destinations or you have any questions about flight arrangements, please contact the workshop organizers for approval before booking the tickets.
2. Accommodations: The Westin Arlington Gateway offers a rate to workshop participants at \$229/night. This is the maximum daily rate that will be reimbursed. Reimbursement requests for other hotels will not be approved. The reservations to obtain the workshop special rate must be made through this link:
<https://www.starwoodmeeting.com/StarGroupsWeb/res?id=1502200257&key=127A4Bo>
B. Room Internet will be free with that group rate. Other incidental expenses will not be covered.
3. Meals: Breakfast and lunch will be provided at the workshop. A workshop dinner will be organized on Thursday evening. Meals (reasonably priced) while on transit are reimbursable but only with itemized receipts. The hotel bill or credit statement is not enough for reimbursement of meals. Meal receipts need to be original and must be in detail showing the individual costs in which food was consumed. Alcohol is not reimbursable and you must note it on the receipt so it can be deducted. If a detailed receipt is not given to you at the time of payment, please ask for a written one with their company name in it and with proof of payment.
4. Ground Transportation: Taxi/Uber to/from both home and Washington airports. Please remember to request a receipt from the taxi driver otherwise you will not be reimbursed. Car rentals will not be reimbursed.
5. Round trip mileage **with prior approval** for personal vehicle use will be reimbursed at .575 cents per mile. You must clearly identify the full address of the trip’s origin and number of miles traveled per trip. A printout of an online mapping direction showing the mileage can be submitted as backup.

Any others expenses are not reimbursable unless pre-approved. For any questions, please send an email to anava@isi.edu with “NSF IIS-GEO reimbursement” in the subject line.

Mail your original receipts to:

USC/ISI
Attn: Alma Nava
4676 Admiralty Way #1001
Marina del Rey, CA 90292

Please specify a complete mailing address (no P.O Box) for mailing your reimbursement check. We suggest that you keep copies of original receipts for your records. Any reimbursement requests received after April 17, 2015 will not be honored. Reimbursements typically take 3-4 weeks to process.