PACI REU Proposal Application

Project Title: Software and Workflow Engineering in the Kepler Project
Relevant URL: http://kepler-project.org
No. of students: 1

Desired Budget: $6,000.00

Description: Kepler is a dynamic cross-project collaboration bringing together several large-scale NSF/ITR projects (GEON, SEEK, ROADNet), the DOE/SciDAC SDM project, and several other projects including EOL (Encyclopedia of Life), Research Surge Enabled by Cyberinfrastructure (Resurgence) and NIH BIRN-CC (planned) to develop an open source scientific workflow system on top of the well-established, extensible Ptolemy II system. Kepler is an emerging system that allows scientists from various domains (bioinformatics, cheminformatics, ecoinformatics, geoinformatics, astrophysics, ...) to design and execute scientific workflows. Scientific workflows can be used to combine data integration, analysis, and visualization steps into larger, automated "knowledge discovery pipelines" and "grid workflows".

There are a number of technical challenges in Kepler, ranging from software engineering and practice issues (distributed development teams, evolving requirements and specifications etc) to current research issues in Grid infrastructure, web service composition, and Semantic Web extensions for workflows.

Therefore, working on Kepler provides students with a unique research experience that includes practical training in real-world tools and best practices (CVS repository, ant, bugzilla, code reviews, ...) as well as the latests developments in cyberinfrastructure research (web service composition, Grid workflows, and semantic extensions).

For this project, the REU student initially will help create a comprehensive software engineering practice for the Kepler project. The student will also act in updating the Kepler website and presentation of the existing work. This will help the student to develop interest on combining his/her computer science skills with different scientific domains and workflows.

After a brief training phase, a specific project will be assigned to the student, e.g., design and implementation of a concrete scientific workflow (domain science centric), or the design and development of a new Kepler component ("actor" or "director"), e.g., to plan Grid workflows.

The project will require good programming skills in Java. Knowledge or experience in one or more of web services, databases, workflows, scripting languages (e.g. Perl, Python), and XML are a plus.

Team Members, Affiliations: Dr. Bertram Ludaescher will interact with the student as the chief scientist for the Kepler project. Dr. Ludaescher acts as a PI/co-PI on some of the participating Kepler projects, and also is an Assoc. Professor at the Dept. of Computer Science and Genome Center, UC Davis, where he is developing new classes in Scientific Data and Workflow Management using Kepler.

Plan to Support W/M, Persons w/disabilities: In order to find a suitable candidate, this position will be publicized to UCSD students in multiple departments, including the Department of Computer Science and Engineering, and the Department of Bioengineering. This will allow us
to reach a very large number of students, many of whom are women, minorities, and/or persons with disabilities.

Two emails of support from project members: 1..
Dear Ilkay,

Your proposal to obtain funding for a student to work with you in the Kepler Project is quite important for our GEON project as well. GEON has benefitted significantly from your work and I am happy to hear that you plan on expanding your workforce with this REU funding. We will be happy to work with you and define tasks that benefit both your research and GEON goals. Good luck and let us know about the outcome.

Best wishes,

Dogan Seber
GEON Project Manager
Geoinformatics Lab
San Diego Supercomputer Center
U.C. San Diego
La Jolla, CA 92093

Plan to integrate students: The REU student will be advised on best software practices like requirements analysis and design, rapid prototyping, and documentation. In addition to the weekly meetings with Mrs. Altintas, the REU student will interact with Dr Ludaescher and the rest of the Kepler team when required and will also participate in weekly lab meetings. In addition the REU student will prepare periodic reports on the progress of the project together with Mrs. Altintas.