Ilkay Altintas

SDSC

# π Person of the Year

## Keeping *the* Flow *in* Scientific Computing

Ilkay Altintas, deputy coordinator for research and director of SDSC's Scientific Workflow Automation Technologies Laboratory, was recently named SDSC's first "Pi Person of the Year." Named after the π symbol, the award recognizes researchers who, as collaborators and innovators in applied research and development, have one 'leg' in one or more science domains, and the other in cyberinfrastructure technologies. Altintas makes sure those researchers can focus their resources on their science rather than solving workflow issues or other computational problems. Altintas has been exploring scientific workflows for over a decade, and is one of the founders of the Kepler collaboration, a project that provides researchers the means to access, arrange, and share data and workflows via a common interface.

### What is a scientific workflow and Kepler in particular?

**Altintas:** A scientific workflow is a series of computational steps that scientists use to generate results. That may involve accessing multiple applications and databases, and processing the data using computationally intensive jobs on high-performance clusters. Kepler is an open source, community-based scientific workflow application that helps users share and reuse data, workflows, and components developed by the scientific community to address common needs. It really started as a grassroots effort just as the compute grid started taking off, and over the past decade has evolved as integrating software has become increasingly difficult. It is named Kepler, after the Ptolemy software on which it is built.

### How do you make the scientific community aware of Kepler?

**Altintas:** In the early days, we had to pitch the tools to communities, go to the researchers, and explain how it would fit in with their work. These days we still recruit, but there is a new generation of scientists that find workflows critical to the success of their work even though they might not be programmers or have the money to hire one. When they become interested in using workflows they find us.

## What science domains use Kepler?

**Altintas:** Kepler has been adopted by researchers in all scientific fields—geophysics, astrophysics, computational biology, ecology, and more. As the need for workflows has increased, I've been involved in projects ranging from bioinformatics and microbial ecology, to ocean monitoring and wildfire management via sensor networks. They all face similar challenges in software and data integration. UC San Diego has a number of biomedical and cancer projects such as the National Biomedical Computation Resource that use Kepler. A new project, called WIFIRE, will apply workflows to wildfire data, which serves not just scientists, but people in the San Diego community who might be affected by such fires (see page 38).

## What do scientific workflows have to do with wildfires?

**Altintas:** San Diego county has one of the most wired environments in the world with the High Performance Wireless Research and Education Network (HPWREN) remote sensing network, and we have frequent wildfires. The WIFIRE project combines the data from those sensors along with other information such as satellite and weather data, and then applies large-scale computing to signal processing, visualization, modeling, and data assimilation to help monitor and predict wildfire behavior. Scientific workflows help simplify those steps so that we can quickly generate the different kinds of models and image data needed by scientists, policy managers, and firefighters, among others.

## What projects would you take on or hardware would you buy if you had unlimited funding?

**Altintas:** It sounds funny to say, but I wouldn't change a thing. I'd try to do what we already do better, more efficiently. Lately I've been interested in "reproducible science." Scientific collaboration just goes forward faster and it is important to capture scientific efforts in a more open way so that it can be reproduced in the future.

## You serve on the editorial board of the *Future Generation Computer Systems* journal. What are popular topics?

**Altintas:** The journal is focused on advanced cyberinfrastructure for large-scale distributed systems, so some of the more popular papers are about workflows, cloud computing, and solving optimization problems.

## How did you wind up at SDSC?

**Altintas:** I had just finished my master's thesis and was working at the Middle East Technical University in Turkey. I saw an interesting job at SDSC working on scientific data management, applied, and wound up moving to San Diego to work as a programmer in 2001. I've been at SDSC ever since, and a lot of good things happened along the way!