



# OMII-UK NEWS

www.omii.ac.uk

December 2006

Inside this issue:

Introducing OMII-UK	1
OMII-UK at Supercomputing 2006	2
OGSA-DAI and GGF18	2
<sup>my</sup> Experiment Portal Party	3
OMII-UK & OMII-Europe	3
Workshop on Workflow Optimisations in Distributed Environments	4
OGSA-DAI activity week	4
Beihang University visits the University of Southampton	5
OGSA HPC Profile Success at Supercomputing 2006	5
Taverna at Indiana University	5
Grid Search for Cancer Prevention: An OGSA-DAI/WebDB application system	6
OMII-UK Software Release — 3.2.0	7
Taverna is moving to Raven	7
OGSA-DAI in Unicore and gLite	8
The OMII-UK Team	8

## INTRODUCING OMII-UK



*The OMII-UK Team - see page 8 for details*

The Open Middleware Infrastructure Institute-UK (OMII-UK) was formed in January 2006 as a collaboration between the School of Electronics and Computer Science at the University of Southampton, the OGSA-DAI project, the Edinburgh Parallel Computing Centre (EPCC), and the <sup>my</sup>Grid project at the School of Computer Science at The University of Manchester.

### Software, support and sustainability

Funded by the **UK e-Science Core Programme**, OMII-UK provides software and support to enable a sustained future for the UK e-Science community and its international collaborators. We support open-source software development by providing a catalogue of e-Science software and a repository for contributing and downloading software – and we invest in community developers to produce the functionality required by our diverse user community.

Drawing upon this software and other packages from the open-source developer community, we provide easy to install and use open-source software releases that provide a secure web service hosting environment, web services and the necessary tools and environments to access these services. This software is supported through comprehensive documentation and training.

### Community-minded

OMII-UK gives confidence to the user community in adopting e-Science solutions. The OMII-UK team now has thirty five members – even more if we include the teams that OMII-UK funds in the open-source e-Science developer community and the many others who take an interest in OMII-UK through its various user communities. OMII-UK staff are also engaged with the international community to define, contribute and disseminate best practice and standards.



open middleware  
infrastructure institute uk  
www.omii.ac.uk



*Supercomputing 2006*

## OMII-UK AT SUPERCOMPUTING 2006

**SuperComputing 2006**, the foremost international conference on high performance computing, networking, storage and analysis was held in Tampa, Florida in November this year. Establishing its presence in 1988, it combines technical and educational programs, workshops and tutorials, in addition to an exhibition hall housing exhibits from both industry and academic research. With over 7,000 registered attendees, and 274 exhibitors spread across a 200,000 square foot exhibition hall, it presents tremendous opportunities for outreach and collaborative activities. OMII-UK was involved in the UK e-Science stand which housed a total of 10 demonstrations for the entirety of the exhibition.

A variety of OMII-UK demonstrators were prepared for the event, with the theme of our primary hosted demos being integration and interoperability. The **HPC Profile Grid Interoperability demonstration** (see page 5) was a new addition and a major success, demonstrating interoperability between a variety of grid architectures and different implementations. It also demonstrated the broad adoption of emerging OGF (Open Grid Forum) standards.

The **Taverna** demo showcased integration of the Taverna client-side workflow environment with **GridSAM** and **Grimoires**, also supplied in our software release. This attracted wide attention and interest from both end users and grid solution providers seeking a means to bridge the interface gap between these two groups.

Also demonstrated was the result of a collaborative endeavour to illustrate **OMII-China's** scheduler service from the **CROWN Grid** software distribution running within the OMII-UK's server infrastructure, successfully interoperating with GridSAM.

A variety of application providers expressed an interest with GridSAM itself, with the objective of using it to facilitate a computational back-end for their applications to take advantage of the grid.

### Further research

Around the exhibition hall there were examples of OMII-UK products being used as the basis for further research. On the AIST stand, an RDF triple store extension to **OGSA-DAI** was demonstrated along with research into "RDF Query Processing

in a Distributed Environment" and "Semantic Grid Resource Discovery and Selection" which both used OGSA-DAI as a framework. Elsewhere around the show, the Grid-Miner project based in Vienna demonstrated a **BPEL**-based workflow enactor which can orchestrate data mining operations using OGSA-DAI.

### A success for UK e-science

With over 700 attendees registering their details on the e-Science booth by the end of the exhibition, the event was deemed a great success for both UK e-Science and OMII-UK.

*Steve Crouch, Hugo Mills and Neil Chue Hong.*

Photo of Supercomputing 2006 courtesy of the Council for the Central Laboratory of the Research Councils.

## OGSA-DAI AND GGF18

The second week in September saw Washington DC play host to both **Globus WORLD 2006** and the **18<sup>th</sup> Global Grid Forum (GGF18)**. For the first time these were run concurrently with **GridWorld 2006**. The Washington Convention Centre provided an air-conditioned oasis within the sticky humidity of DC. A half dozen members of the OGSA-DAI team attended the conference. There, we hosted a developers session and a user forum and also contributed to a session on a Globus data management tool. The new de-

sign of OGSA-DAI, planned for release in March 2007 was presented (see "OGSA-DAI Activity Week" on page 4 for details on this new design). This design includes a completely refactored activity framework, an expanded resource model, improved security and support for persistence, scalability and robustness. The proposal met with a positive reception. We also had a productive meeting with our Globus friends to discuss architectural issues.

*Mike Jackson.*



**OGSA-DAI**

"We also had a productive meeting with our Globus friends..."

## myEXPERIMENT PORTAL PARTY

myExperiment – an exciting new solution which supports the experimental process for the e-researcher – took an important step forward when the myGrid team held the myExperiment Portal Party on 28th & 29th September, with hand-picked users and portal developers from around Europe attending to help us determine priorities.



*The Portal Party*

If you have any questions  
please contact  
support@omii.ac.uk

The Portal Party taught us that our users want to share and discover workflows in an easy-to-use environment with a sense of community. Another key driver was the ability to launch workflows directly from that environment without ever having to "see" Taverna. The portal will enable the running of workflows from the web and also act as a workflow repository, enabling the sharing of workflows and resources between myGrid users. We are now in the process of determining the set of features required in the first release. Please see the Portal Party wiki for details [www.mygrid.org.uk/wiki/Portal/](http://www.mygrid.org.uk/wiki/Portal/).

*June Finch.*

## OMII-UK & OMII-EUROPE

OMII-Europe ([www.omii-europe.org](http://www.omii-europe.org)) is an EU Framework 6 funded project consisting of 16 partners from across Europe and beyond. Its goal is to achieve standards driven service interoperability across:

- Computational job submission using the **Job Submission Description Language (JSDL)** and the **Basic Execution Service (BES)**
- Data access and federation using OGSA-DAI. This software will soon be ported to work within UNICORE and gLite (more info later)
- Accounting using the **Usage Record (UR)** and **Resource Usage Service (RUS)** with extensions proposed by OMII-Europe
- The use of **Virtual Organization Membership Service (VOMS)** as a tool for managing virtual organisations

These services will operate across the major Grid Middleware stacks in use within the EU — EGEE's gLite, Unicore, GT4 from the Globus Alliance, OMII-UK, and CROWN Grid. In addition, a common security infrastructure will be developed and a portal, based around GridSphere,

used to access these services. All software deliverables will be held in a repository, and an automated build and test system used to provide binary builds across multiple platforms which are verified for their correctness and compliance with relevant standards.

OMII-UK partners at Edinburgh and Southampton are contributing to work packages within the OMII-Europe consortium. This involves:

- Support process activity linking OMII-Europe and OMII-UK helpdesks and documentation
- Build and Test environment involving collaborations with the **ETICS** project at CERN and the Condor team at the University of Madison at Wisconsin
- Job Submission activity and interchange of OMII-UK and **CROWN Grid** software components

- Knowledge transfer between OMII-UK and OMII-Europe repository activities

On the managerial level the Director of OMII-UK Steven Newhouse chairs the OMII-Europe Technical Committee and contributes to the Repository activity and many of the other work packages as technical coordinator.

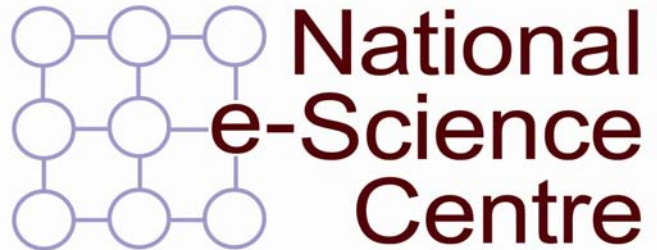
EPCC is a full partner of OMII-Europe and OMII-UK. They have representation on the Project Board, Management and Technical Committees. They lead the SA3 support activity which provides support for OMII-Europe developed components deployed on an Evaluation Infrastructure. In addition they lead the JRA1 Dataservices re-engineering activity which is making OGSA-DAI available under the UNICORE and gLite middleware stacks. The JRA1 Dataservices activity will be able to utilise the expertise of the OGSA-DAI development team who are also based at EPCC.

*Tim Parkinson and Steven Newhouse.*



## WORKSHOP ON WORKFLOW OPTIMISATIONS IN DISTRIBUTED ENVIRONMENTS

This workshop on workflow optimisation was held at the e-Science Institute at Edinburgh this October. It brought together a number of researchers and practitioners from academia and commerce interested in the area of workflow optimisation for scientific applications. A wide range of workflow optimisation talks were presented at different levels of implementation. At a lower level, workflow scheduling optimisation techniques for Directed Acyclic Graph representations were presented together with some good example use cases.



Examples of time and space optimisations of scientific workflows in distributed environments using the **Pegasus** system were demonstrated with some interesting real world application scenarios from the National Virtual Observatory in U.S. and the Southern California Earthquake Center. Higher levels of workflow optimisation were demonstrated achieving performance gains by being able to dynamically select a service based on the load and CPU speed of the

host on which the service was deployed.

The workflow used by **OGSA-DAI** was presented and the key differences from the more traditional workflow systems were given. The new OGSA-DAI enactment model, targeted for the next release, was described in some detail in order to pinpoint the optimisation improvements achieved in the pipelining processing that is used by OGSA-

DAI. Finally, at a higher-level, workflow restructuring was discussed regarding improving the performance of OGSA-DAI.

Powerpoint presentations can be found online at:

[www.nesc.ac.uk/action/esi/contribution.cfm?Title=702](http://www.nesc.ac.uk/action/esi/contribution.cfm?Title=702)

*Kostas Karasavvas.*

## OGSA-DAI ACTIVITY WEEK

The OGSA-DAI development team is currently working on a major new version of the project scheduled for release in spring 2007. OGSA-DAI allows users to construct workflows that chain together a number of activities. These activities implement various data operations such as querying a database, transforming data or delivering data to remote systems.

One of the goals driving the design of the next release is to produce a new activity model that allows for better activity reuse and interoperability. During the first week of Octo-

ber the OGSA-DAI development team worked intensively in a highly collaborative and agile manner to implement the activities required for eight new scenarios that OGSA-DAI has previously been unable to support. All the scenarios were implemented successfully and have largely validated the design decisions that will form the basis of the new release.

It is hoped to arrange similar intensive development weeks in the future on other issues such as security or data federation. We would be espe-

cially keen to include a few interested parties from outside the core OGSA-DAI development team in these future exercises. Please get in touch if you may be interested.

A document describing the scenarios of the activity week as well as design documents relating to the ongoing OGSA-DAI development are available from: [www.ogsadai.org.uk/documentation/Design\\_documents](http://www.ogsadai.org.uk/documentation/Design_documents).

*Ally Hume.*

## BEIHANG UNIVERSITY VISITS THE UNIVERSITY OF SOUTHAMPTON



*OMII-UK and Beihang University collaboration*

On the 9<sup>th</sup> Oct 2006 a delegation from one of China's leading scientific Universities visited OMII-UK Southampton. They were welcomed by Prof Peter Henderson. Prof Jin-Peng Huai (Chairman of the Institute for Advanced Computing Technology) and Prof DianFu Ma (Dean of Graduate Education) gave a presentation that demonstrated Beihang's pre-eminence in Computer Science within China. Steven Newhouse in-

troduced the OMII-UK organisation and Alistair Dunlop described the OMII-Europe project. While the senior staff discussed inter-campus collaboration, their three technical staff met with Stephen Crouch and Gary Li to discuss job submission interoperability between OMII-UK and CROWN Grid

*Tim Parkinson.*

## OGSA HPC PROFILE SUCCESS AT SUPERCOMPUTING 2006

OMII-UK is involved in both the Open Grid Services Architecture (OGSA) working group and many of the OGSA sub groups within the Open Grid Forum (OGF), which have strong representation from both industry and academia.

One important standard being developed by OGSA is the **Basic Execution Service (BES)**. This proposed standard will form the basis for interoperability between different implementations of Grid middleware, providing a common interface for submission of jobs – using the **Job Submission Description Language (JDSL)** – to compute resources. The use of the BES specification is being further defined within the OGSA

**HPC Profile working group.** The HPC Profile of OGSA BES has, over the last few months, been implemented by ten different organisations, including Microsoft, Fujitsu, an OMII-UK project – GridSAM, and the CROWN project from OMII-China. The interop demonstration at this year's Supercomputing involved each of the projects with an HPC-Profile implementation showing their client software submitting jobs to each of the other implementations' endpoints.

What made this demonstration remarkable was that it involved so many participant organisations spread across the exhibition hall, each demonstrating their ability to

interoperate with the implementations of the other participants. By the end of the show, every project had managed to interoperate with every other one, with only a few problems remaining unsolved. Whilst it may seem like a small victory, demonstrating a limited form of a provisional standard for one small part of a Grid infrastructure, this interop demonstration is actually a major step, as for the first time it means that a wide variety of different Grid middleware products can interoperate using a common interface. This is a strong vote of confidence from all of the participants in the work of the OGF and OGSA, and shows the degree of engagement from the major developers in the field, including OMII-UK and Globus.

*Hugo Mills and Steve Crouch.*

## TAVERNA AT INDIANA UNIVERSITY

The Chemical Informatics and Cyberinfrastructure Collaboratory (CICC) project is funded by the NIH (National Institutes of Health) at the University of Indiana. The project focuses on developing workflows which encapsulate important processes in chemoinformatics and drug design, use diverse kinds of informa-

tion together in novel ways, and which are of demonstrated scientific merit. To achieve its goals the project is successfully using Taverna for workflows and services management.

The influence of Taverna on the project is such that throughout the pro-

ject's website ([www.chembiogrid.org/wiki/index.php/Workflows](http://www.chembiogrid.org/wiki/index.php/Workflows)) there are examples of workflows built using Taverna, "getting started" tutorial and movie of workflow running in Taverna. The project Manager, Marlon Pierce, described Taverna as an excellent open source workflow engine.

*Franck Tanoh.*

# GRID SEARCH FOR CANCER PREVENTION: AN OGSA-DAI/WEBDB APPLICATION SYSTEM

**Business Search Technologies (BST)** Corporation is a private company based in Tokyo. Set up in February 2004, it has been working on software research and development over three technical areas: search technology, internationalization (I18N), and grid computing. BST has developed a security enhanced and multi-lingual capable commercial search engine package called WiSE (Worldwide Internet & Intranet Search Engine), and a software internationalisation tool, called "World Wide Navi".

In the grid computing area, BST has been working closely with **Grid Technology Research Center/ National Institute of Advanced Industrial Science and Technology (GTRC/AIST)**, a Japanese governmental research institute based in Tsukuba and Tokyo. BST has been involved in two joint research efforts with GTRC: GridASP (see [www.gridasp.org](http://www.gridasp.org) for details) and the data grid. In November of 2005, BST was contacted by CBIE (Clinical and Bio-Informatic Engineering), a division of the Graduate School of Medicine at the University of Tokyo, and asked if they would be interested in working on a project, using OGSA-WebDB in the fight against cancer. The researchers were looking for a way to integrate information from various databases, including public web sites, in the search for a way to prevent cancer.

## A New Partnership

Having agreed that it would be keen to get involved, BST joined a project team in partnership with CBIE, and with GTRC/AIST. The aim of this project was to develop a secure, integrated query interface to medical information spread over multiple databases, both public and private. These databases would include the JSNP (Japanese Single Nucleotide Polymorphism), a database of common gene variations in the Japanese population, OMIM (Online Mendelian Inheritance in Man), a database of human genes and genetic disorders developed by staff at the Johns Hopkins University in Baltimore Maryland, PubMed, a system developed by the National Center for Biotechnology Information (NCBI) at the US National Institute for Health to offer access to citations from biomedical literature, and PharmGKB, a pharmacogenetics and pharmacogenomics knowledge base.

## Technologies Used

To do this, the team used two main technologies: OGSA-DAI middleware, to grid-enable the data resources, and OGSA-WebDB, to integrate the web databases with the OGSA-DAI environment.

"We chose to use the OGSA-DAI technologies as the basis of our OGSA-WebDB, because it is an open software, and also because we knew it could help us to retrieve the data we wanted from the web sites," said Isao Kojima, Senior Research Scientist, Leader of the Data Grid Team at GTRC. "I worked closely with the OGSA-DAI team in Edinburgh, and had a lot of communication with them," mentioned Mirza Phalevi Said, Research Staff of Data Grid Team at GTRC/AIST. "They were very kind, and gave us a lot of advice."

"We positioned OGSA-DAI and OGSA-WebDB as the framework for building GSCP," said Jun-ichi Okamura, the architect of GSCP system from BST.

"As far as I know, coupling of OGSA-DAI and OGSA-WebDB is the only technology available that can provide access to all the web and relational databases we wanted to access, and then allows you to process the data on the grid."

## Results

A single query on the system developed by the team was able to access all of the databases, and the results would then be combined to show up on a single GridSearch screen, so that they could all be viewed as if they were coming from local data sources. Wrappers on the four databases being accessed inserted the extracted data into proxy databases,

which were then accessed via the OGSA-WebDB and OGSA-DAI interfaces. Integration of outside web databases in such a manner has revealed an important feedback to the team: "We were dealing with outside databases, outside our control. And some of those have long response times, and there's nothing you can do about that. We had to enhance WebDB so that it can let us control the access time as needed." Thus, WebDB was improved to have several adjustable parameters such as, "connection retry number" and "maximum number of query results".

The first version of Grid Search for Cancer Prevention, V1.0, was released in March of this year. In future, the team hopes to be able to access more resources, with secure access to private data, and also more integration at the syntax level.

"One issue is that from time to time the outside sites change their interface – and that's beyond our control, we can't prevent it. So an end user of our system discovered one day that queries were failing, and he couldn't imagine why. So he called us up and we investigated, and that's when we realised it had changed. Currently we don't have an automated method of following and spotting that sort of change – we would have to call the managers of the sites to ask about their plans. "Currently we don't really have a good idea of how to automate the process, but that's what we want to do", said Ryuichi Yoshida, a manager at BST's Research and Development Centre.

*Gillian Law.*

## OMII-UK SOFTWARE RELEASE - 3.2.0

The latest versions of the OMII-UK Server & Client software releases are now available for download from the website (under the Downloads menu).

These releases install on a variety of modern Linux distributions, Windows XP SP2 (client only) and Apple OSX. They will run in a variety of Java JDK environments. Documentation (Installation & Setup Guide, User Guide and Training Guide) are available from [www.omii.ac.uk](http://www.omii.ac.uk). This release is a fully functional production release.

Several new software components have been integrated into this release:

- **Taverna workflow tool** — Taverna (v1.4.1)
- **BPEL Workflow Editor and Execution Engine** (v2.1.0)
- **Application Hosting Environment** — (v1.0.2)

The Taverna workflow environment from the <sup>my</sup>Grid project supports the definition of workflows between secure web services using X.509 certificate based authentication within WS-Security. The BPEL Workflow Editor uses an **Eclipse** plugin to provide an environment for the definition of workflows using the Business Process Execution Language (BPEL). The **ActiveBPEL** execution engine is used to run these workflows and uses X.509 based WS-Security for workflow submission and execution. The Application Hosting Environment (AHE), and the pre-requisite WSRF::Lite component, are available on a subset of our supported platforms (32-bit Linux).

A refactored server installation that downloads the individual components on demand will enable us to expand

the number of supported platforms for these components in later releases.

The following software components are included in this release (new in 3.1.0):

- Job submission and monitoring web service — GridSAM (v2.0.0)
- WS-Reliability and WS-ReliableMessaging implementation — FIRMS (v1.2.0)
- WS-Eventing implementation — FINS (v1.2.0)

GridSAM 2.0.0, conforms to the recently approved Job Submission Description Language (JSDL) standard from the Open Grid Forum and has an improved authorisation mechanism. The documentation and installation speed of FINS and FIRMS have been improved significantly. Source code for the FINS & FIRMS projects is now available from NeSCForge.

The documentation for this release is available online and as a separate download from the repository (in HTML and PDF versions). There are three guides: Installation, User Guide and Training Guide. The Training Guide has been enhanced to provide tutorials on GridSAM, Account Service, BPEL and the AHE.

This release uses Tomcat 5.0.25, Axis 1.2.1 and a patched version of WSS4J to provide the secure Web Services infrastructure. Instructions are provided on how you can use an existing database instance instead of the PostgreSQL instance provided in the server release.

*Steven Newhouse.*

## TAVERNA IS MOVING TO RAVEN



**Raven** is a framework for distributing and launching modular Java applications.

Based on the popular build system **Maven**, Raven resolves dependencies and downloads any missing or updated .jar files. Combined with SPI-technology for

module discovery, this allows it to distribute a self-updating and extendable Taverna. For instance, if the SoapLab processor needs to be updated due to changes at the service provider, we can release a new version of the SoapLab module, which can then be automatically picked up by "older" Taverna installations. This can be compared to package distribution systems for Linux. Using Raven changes the Taverna distribution from just being a set of .jar files to Raven with a Raven-profile, which specifies Taverna and its dependencies.

*Stian Soiland.*

## OGSA-DAI IN UNICORE AND GLITE

The OGSA-DAI software is going to be ported to work within the Unicore (Unicore 6) and EGEE (gLite 3) middleware stacks thanks to the OMII-Europe project. The intent is that alpha versions of the OGSA-DAI software will work within these environments by May 2007 and full release quality software will be made available by May 2008. OGSA-DAI can already be deployed on to GT4,

the OMII-UK container and vanilla versions of Axis on Tomcat. This will allow the data access and integration capabilities offered by OGSA-DAI to be exploited within these new environments and expand the user base of OGSA-DAI. More information about OMII-Europe can be found at <http://www.omii-europe.org>.

*Mario Antonioletti.*

“The intent is that alpha versions of the OGSA-DAI software will work within these environments by May 2007 and full release quality software will be made available by May 2008.”

## THE OMII-UK TEAM



From left-to right: Mario Antonioletti, Bryan Carpenter, Gary Li, Jodi Crisp, Hugo Mills, Steven Newhouse, Claire Walker, Juri Papay, David De Roure, Victor Chang, Muan Hong Ng, Tim Parkinson, Franck Tanoh, Malcolm Illingworth, Christopher Brown, June Finch, Stephen Crouch, Neil Chue Hong, Ally Hume, Stian Soiland, Justin Bradley, Richard Boardman, Malcolm Atkinson, Amy Krause, Mike Jackson, Carole Goble, Elias Theocharopoulos, Stuart Owen, Kostas Karasavvas, Tom Sugden, Bartosz Dobrzelecki, Daniele Turi, David Withers, Katy Wolstencroft, Tom Oinn.

## CLOSING WORDS

We hope that you have enjoyed this issue of the OMII-UK newsletter. We would love to hear what you think and any suggestions as to how this publication can be improved. If you have suggestions for articles that should be included or things that you would like to hear about then please do get in touch. Email: [support@omii.ac.uk](mailto:support@omii.ac.uk) and we will try to take this into account for our next issue.

The Editorial Team (Mario Antonioletti, Juri Papay and Franck Tanoh)