

Sakai Evaluation Exercise
Rob Crouchley and Adrian Fish
e-Social Science Centre of Excellence, University of Lancaster
Rob Allan and Dharmesh Chohan
CCLRC e-Science Centre, Daresbury Laboratory

Executive Summary for JISC VRE Bid Evaluation Panel.

Sakai (<http://www.sakaiproject.org/>) is in our view now the leading example of a Web-based research collaboration framework written in Java and built upon the open source, JSR-168 compliant, uPortal hosting environment. Sakai itself is open standards and open source, and has evolved from the CHEF *Comprehensive Collaborative Framework* suite of tools from University of Michigan, which is widely used in the USA for virtual learning and collaboration. Sakai is standards based and leverages the Open Knowledge Initiative's Open Service Interface Definitions (OSIDs, see <http://web.mit.edu/oki/>), which was developed with funding from the Andrew Mellon Foundation and can be extended for a variety of purposes.

The JISC JCSR-funded *Sakai Evaluation Exercise* has addressed the following areas which are summarised below and will be described in a full report at <http://tardis.dl.ac.uk/ReDRESS/sakai.pdf>:

1. Comparing Sakai with Alternative Frameworks for a VRE;
2. Assessing the Ease of Administration (EoA) of Sakai for a VRE;
3. Establishing the feasibility of making existing VRE (Grid) components available in Sakai;
4. Establishing the issues involved in extending the functionality of Sakai particularly to use Web services for distributed development and deployment;
5. Developing a *Roadmap for a UK Virtual Research Environment*. There is now in a separate report from the JCSR VRE Working Group as cited in the CFP notice 04/05.

As part of the *Sakai Evaluation Exercise*, JISC funded the UK membership of the Sakai Educational Partners' Programme (SEPP) which enabled the authors to attend the first developers' conference, Denver, 23-25/7/04. Some of the comments below are based on discussions with developers in Denver and in prior and subsequent exchanges about Sakai in a Service Oriented Architecture (SOA) VRE or VLE. Other institutional members of SEPP include University of Cambridge and University of Hull.

As part of this evaluation, we have recently put up a simplified Sakai demonstrator at <http://redress.lancs.ac.uk/sakai-uPortal> (username=guest, passwd=eResearch). This will allow you to view the initial version of the proposed Sakai Evaluation worksite. A prototype portal for the National Grid Service is also available at <http://thames.dl.ac.uk:8080/chef/portal> (username=guest, passwd=eResearch). This is based on the OGCE portal (using CHEF) and its Grid tools will be ported to Sakai in collaboration with the TeraGrid developers at Universities of Indiana and Texas.

Comparing Sakai with alternative Frameworks for a VRE.

This part of the work focused on Web-based access to (mostly remote) VRE services. We wanted to ensure that service and resources already funded by JISC could be fully integrated. Many of the collaboration tools in Sakai run locally on the Web server, but greatly enhance the VRE for a working scientist. Other tools which might be components of a VRE include drag 'n' drop desktop interfaces, programming libraries and GUIs, but Web clients are more familiar and therefore extremely popular.

Early Grid portals included HotPage and GridPort from San Diego Supercomputer Centre, Grid Resource Broker from University of Lecce, Italy, HPCPortal and InfoPortal from the Daresbury Group. Developers from SDSC have mostly moved to Austin in Texas. More recently the GPKD toolkit from NCSA has been used in Java-based portals, e.g. for the UK DAME project and the US Alliance portal. Together these gave rise to the US Portal Project and the focus of the GGF Grid Computing Environments working group whose members produced the CHEF framework and OGCE *Open Grid Computing Environment* portal now adopted by the US NMI *National Middleware Initiative* for many leading e-Science laboratories such as NEESGrid and CMCS. Sakai is the evolution of CHEF for a Collaborative Learning Environment (CLE) which will soon also contain the Grid tools being ported by the TeraGrid developers at Universities of Indiana and Texas.

The only comparable alternative is GridSphere from the European GridLab project, see <http://tardis.dl.ac.uk/Papers/Portals/portals.pdf>. This is also based on JSR-168 standard portlets, but has less collaboration tools, no e-Learning tools and supports a much smaller Grid community which does however include the UK RealityGrid and GeneGrid projects. Portlets have been shown to easily plug into either framework. A number of other Web-based frameworks exist for specific purposes, such as Bodington, but do not provide a standard-compliant interface for community contribution of e-Research tools. Peer to peer tools are now growing in popularity but are not yet sufficiently mature to be used in a large-scale distributed VRE. Products such as LionShare should however be integrated into a user-facing portal via an

appropriate interface.

We conclude that Sakai is the leading 2nd generation Web portal for both e-learning and e-Research with the backing of a large international user community

Assessing the ease of Administration (EoA) of Sakai for a VRE.

Several instances of CHEF, Sakai RC1 and Sakai RC2 have now been deployed for ReDRess: <http://redress.lancs.ac.uk/sakai-uPortal> and for HPCPortal v3.0 targeting the National Grid Service: <http://thames.dl.ac.uk:8080/chef/portal>. These have proven to be flexible and easy to administer with several project work sites and a growing number of developers and users keen to use the collaboration tools.

Establishing the Feasibility of making existing VRE (Grid) Components available in Sakai.

During the evaluation work, native components have been ported into the Sakai framework. We stress that the OKI interface has not yet been used in this work as it focuses mainly on issues of managed learning. We are however documenting issues defining an "integration API" which could eventually extend the OSIDs. Areas for framework extension which have been discussed at length with the Sakai and OGCE developers and presented in the JISC Roadmap for a UK Virtual Research Environment include: a full policy-based security model; service discovery and invocation; description of a collaborative research session; link to peer-to-peer systems and group management. Components which have been imported into Sakai during this evaluation can be seen via the URLs mentioned above and include:

1. Authentication via organisational LDAP, though you can't see this as its hidden from the user;
2. CopperCore learning content player, Sakai VLE site at Lancaster;
3. InfoPortal static and dynamic data services, Sakai VRE site at Lancaster;
4. OGCE Grid tools linking to the JCSR-funded National Grid Service clusters, CHEFsite at Daresbury;
5. RSS news feeds for ReDRess and NGS, feed from Daresbury to both Sakai and CHEF sites;
6. IRC, Internet Relay Chat tool, Sakai VRE site at Lancaster;
7. CLAWS, a linguistic tool for tagging text, Sakai VRE site at Lancaster.

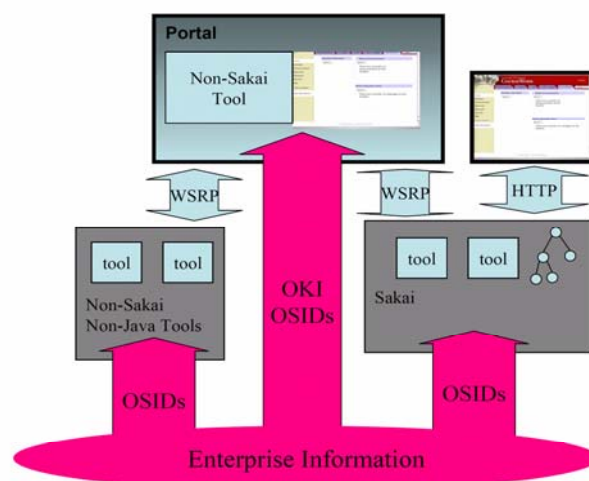
Further details on these components are to be found in Appendix 1.

Establishing the Issues involved in extending the Functionality of Sakai.

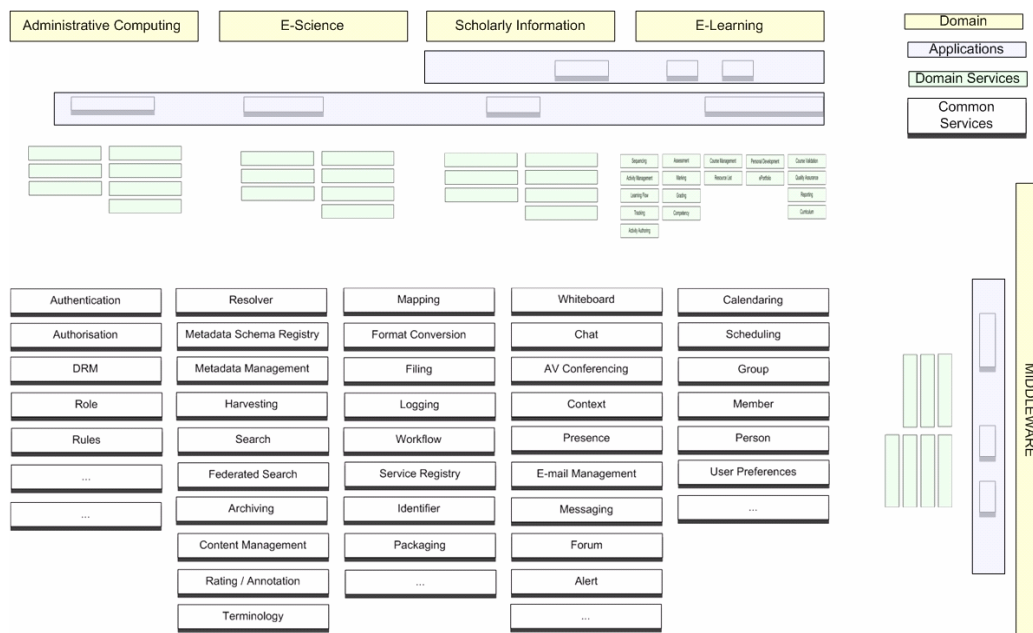
The "Integration API" was also mentioned above. The background has been presented in the VRE Roadmap. These issues are important in enabling the full spectrum of VRE components to be presented in any framework, whether or not Web based. We believe that a portal provides sufficient encapsulation and management capabilities that most of the issues can be addressed, possibly with the exception of complex scientific visualisation which can however be enabled via Java WebStart or similar download tools.

We have had many exchanges about WSRP (via reviews, proposals, face to face meetings, emails) with the Sakai developers over the last 9 months. We think these exchanges have been very successful. The Sakai Chief Architect Chuck Severance has just obtained agreement from the Sakai Board for a major shift to WSRP-based architecture, see below.

Integration Architecture



For full details on the case made see http://www.immagic.com/eLibrary/TECH/U_MICH/S040826S.pdf. This move will make Sakai an ideal hosting framework for any combination of tools and services from the e-learning, e-Science, the information environment and the institutional strands of activities shown in the Cross Domain Modelling and Common Services Figure (see below) from the recent paper titled *Service-Oriented Frameworks: Modelling the infrastructure for the next generation of e-Learning Systems*. DEST (Australia), JISC-CETIS (UK), and Industry Canada.



Appendix 1. Sakai Evaluation: VRE Demonstrator

As part of our evaluation, we set up a simplified Sakai VRE demonstrator at <http://redress.lancs.ac.uk/sakai-uPortal> (username=guest, passwd=eResearch) in order to establish the difficulties of porting tools to Sakai. On the **Sakai Evaluation** worksite and in addition to standard Sakai collaboration tools (discussion, chat etc) there are the following components:

- **IRC Chat Tool.** IRC is a widely used protocol for on-line chat sessions including group or 1 on 1. Users can choose between several threads. A Chat portlet is already present in the Sakai framework but work is needed to integrate chat sessions from project worksites into a user's personal worksite and to link to external IRC channels. (Jabber is an alternative.)
- **Infoportal.** InfoPortal is a "Grid community portal" which provides access to information tools for Grid resources and services. InfoPortal provides access to two types of data: static (e.g. architectural information) via an XML database; and dynamic (e.g. current and historical usage information) via Globus MDS. It also provides two kinds of views on this data: resource centric (based on full machine name) and site centric (based on geographical location via an "active map"). Links to additional monitoring tools are also provided, e.g. for network monitoring via GridMon. InfoPortal is in many ways complementary to R-GMA which is used by the particle physics community.
- **CLAWS.** CLAWS (Constituent Likelihood Automatic Word-tagging System) is a tool used by linguists, text is passed to the tagging program which assigns a part-of-speech tag to each word or word combination in the text. Because one orthographic form may have several possible parts of speech (e.g. love can be a verb or a noun), after the initial assignment of possible parts of speech to the words in the text, CLAWS uses a probability matrix derived from large bodies of tagged and manually corrected texts to disambiguate the words in the text. The matrix specifies transition probabilities between adjacent tags, for example given that x is an adjective, what is the probability that the item to its immediate right is a noun? Again, these probabilities are constantly updated from new data. CLAWS tracks through each sentence in turn applying these probabilities. Finally, manual post-editing using a special tag editor may take place if desired to correct fully the machine output. The CLAWS system enjoys a success rate in the region of 96%-97% on written texts, and is also successful, though to a slightly lesser degree, on spoken texts.