

A Collaborative Orthopaedic Research Environment.

S Grange[¥] L Gilbert[†], and G Wills[†]

[†] School of Electronics and Computer
Science,
University of Southampton,
SO17 1BJ, UK
+44 (0)23 8059 2831
{lg3, gbw}@ecs.soton.ac.uk

[¥] Royal College of Surgeons of England,
London, UK
bonesculptor@hotmail.com

ABSTRACT

The role of collaboration in scientific and scholarly research is changing due to advances in Web technology. In particular, the need for collaborative science has generated demands for working environments that facilitate human communication and resource sharing among research communities. The Collaborative Orthopaedic Research Environment (CORE) project provides an infrastructure that combines clinical, educational and research activities in a Virtual Research Environment (VRE) for orthopaedic researchers to collaborate in the design, analysis, and dissemination of experiments. An overview of Service-Oriented Architecture (SOA) concepts is presented in this report before moving on to discuss the benefits and rationale of using a SOA in the context of the CORE project. A user requirements study conducted to guide the authors in designing the CORE VRE is also reported.

1. Introduction

The accepted role of scientific and scholarly publication is to record research activity in a timely fashion, keeping others in the research community up to date with current developments. Until very recently, printed journals were the most efficient method for the dissemination and archival of research results. Technical advances in the past decade have allowed the process of scholarly communication to take other forms, particularly in the dissemination and storage of articles via the World Wide Web (Carr *et al*, 2004). In addition advances in Web technology are changing the role of collaboration in scientific and scholarly research. In particular, the need for collaborative science has generated demands for working environments that facilitate human communication and resource sharing among research communities. Such environments have been typically implemented as monolithic systems in the past, which are then faced with challenges in adapting to changing requirements and changing technology.

A service approach leads to more loosely coupled systems, where individual parts may be developed by different people or organizations. Due to the nature of the loose coupling in a SOA, applications can be developed and deployed incrementally. In addition, new features can be easily added after the system is deployed, and existing features can be re-implemented to take advantage of developments in hardware or software. This modularity and extensibility make SOA especially suitable as a platform for collaborative research environment in which there are evolving requirements and standards. Services are also appealing in terms of their ability to be reused, as they have well-defined public interfaces.

CORE is a Collaborative Orthopaedic Research Environment, with a SOA that provides an infrastructure that combines clinical, educational and research activities in a Virtual Research Environment (VRE). This integrated computer support across the research and educational cycles is necessary because these activities are intrinsically coupled and are part of the requirements of the surgeon's Continuing

Professional Development (CPD). CPD is an essential part of the healthcare professions and the use of ICT provides an opportunity to improve the efficiency of both the teaching of and the learning by lifelong learners. Surgeons must make time in their busy schedules to undertake research and publish papers in order to achieve goals under their learning contracts with their Professional Colleges. Mokola (2006) has also shown that it is important in the South African CPD information system for people to have a record of work they have undertaken.

The CORE project allows surgeons to: create technical material (including non-research material for education), analyse data (from their own trials or data entered from journals), investigate hypotheses (from their own work or as meta or thematic reviews), discuss the findings from their own or others work, and prepare and submit articles for review.

This paper will focus on a specific architecture, the CORE VRE, which supports a critical subset of the e-Science cycle: the collation and analysis of experimental results, the organisation of internal project discussions, and the production of appropriate articles depending upon the requirements of conferences and journals. The CORE project aims to provide a set of generic services for conducting research, with a portal framework that is applicable across surgical and medical training.

2. Service-Oriented Architecture

This section presents the SOA concepts used as guidelines in designing and building the CORE VRE using Web services technology.

An SOA can also be viewed as an attempt to modularize large complex systems in such a way as they are composed of independent software components that offer services to one another through well-defined interfaces. These flexible granular functional components expose service behaviours accessible to other applications via Web Services. The service approach is ideally suited to more loosely coupled systems. Wilson, Blinco, & Rehak (2004) discuss in detail the advantages of using SOA:

- *Modularity*: As appropriate services are dynamically coupled, it is relatively

easy to integrate new services into the framework.

- *Interoperability*: Due to standardization of the communication and description of the services, third party services can easily be incorporated as required.
- *Extensibility*: Due to the relative ease with which services can be incorporated into a system, there is less danger of technology 'lock-in'.

Designing Web Services is an example of the more general case of designing for Service Oriented Architecture (SOAs). These SOAs present particular challenges to developers. Dijkman & Dumas (2004) explain the need for particular Service Oriented Design strategies, based on a number of characteristics that differentiate Service from Component-based design :

- *High Autonomy*: Services will likely be developed by distributed autonomous teams.
- *Coarse Granularity*: Services will tend to be more coarse-grained than traditional components, and there is a need to reconcile many aspects of a complex activity.
- *Process Awareness*: Services will be driven by explicit business processes, which mean there is a high relation to enterprise design.

Enterprise-level service development is most affected by the later two characteristics. For example, Quartel, describe the use of design milestones to help develop Web services from business practices (Quartel, *et al* 2004). Other developers have created environments to ease the creation of composite services (Wilson, *et al* 2004). Martin suggest that the best way to implement Web Services in an enterprise is to start with a component-based architecture that exposes business process level services as Web services (Martin, *et al* 2003). This is no different for the research domain where it is essential to start with services that are integral to the task at hand.

3. The User Requirements Study

The CORE VRE focuses on two main areas. The first is the collection of data (operations and clinical trials) from hospital based orthopaedics. The second is a Virtual Bone Laboratory, incorporating the basic science disciplines of molecular and

tissue biology, engineering, and computing, with the allied medical disciplines of pharmacology, prosthetics, trials management, and the clinical disciplines concentrating upon musculoskeletal applications in rheumatology and orthopaedics.

The extent to which specific users can achieve specified goals with effectiveness, efficiency and satisfaction using a VRE is largely driven by their requirements. Hence, the CORE project conducted a study to identify the issues and requirements for the provision of Grid/Web services that relate to the storage, access, use, reuse and dissemination of research data in repositories information in digital libraries under a VRE. The participants in this study comprised both experts in computer science, and end users who were researchers and clinicians working in an orthopaedic (bone) laboratory.

The study was conducted through a consultation process that involved semi-structured interviews and online survey. The end users in the CORE project typify e-scientists and e-learners in their research environment. Semi-structured interviews were conducted with five professionals. The participants were chosen based on their expert knowledge in the research areas of digital libraries, information and learning technologies, and the biomedical domain.

In addition an online survey was conducted in order to identify end user requirements. A survey questionnaire containing nine questions was designed to be completed by participants via email. A total of 17 orthopaedic surgeons (including 14 Higher Surgical Trainees (HST), one consultant, and two surgeons who did not disclose their grade) took part in this survey. HSTs are qualified surgeons (Registrars) training to be consultants. Their study is work-based, and they are rarely co-located. During the six years of training, they usually move post twelve times, and they keep a logbook (e-Portfolio) as required by their learning contracts with their professional Colleges. There are approximately 200 HSTs in the UK at any one time.

Participants who took part in the online survey responded to questions in a way that indicated that they were technologically competent, in that they used computers at home and work, and they discovered and located research materials using the Web.

HSTs have to conduct research and publish papers as part of their training. Hence, all of the participants (except one) were involved in research projects, where 13 of them conducted individual audits/projects for publication, and 3 others undertook a part-time postgraduate degree in conjunction with their training.

3.1 Summary of Findings and Recommendations

In general, the participants were very positive about the CORE VRE and provided many valuable suggestions as to how to enhance the planned functions in the VRE. Sharing of research resources, *i.e.* scientific data and publications, appeared to be an attractive idea. This was supported by the comments given by survey respondents that expressed how time-consuming it was to conduct collaborative research without a space to share scientific data and results.

The results from the study reinforced the approach of developing a Grid/Web service-based research environment, equipped with the ability to evolve with the changing requirements of its user community. For example, there are many protocols to choose from when conducting a clinical trial. Although it might be possible to make a small number of generic trial protocols or templates, it is impractical to predefine and encode every possible experimental procedure. Likewise, it is also difficult to predict every type of data that can be collated and analyzed, and every possible dissemination route that users will follow to publish their results to the wider community. Therefore a loosely coupled architecture is essential to allow the flexibility of adding extra services at a later stage when the user requirements change.

A clear priority requirement arising from the study was to provide users with a single point of entry for a VRE, which would bring together resources (*i.e.* scientific data and publications) from different channels, for their convenience. The study participants also pointed out that the VRE should be made easy to use for those who may not be particularly computer literate. Another important finding from the user requirement study was that different types of users would require data to be presented differently. For instance, biologists are generally interested in graphical representations of data while

mathematicians might prefer numerical presentation of the same set of data.

The authors also learned from the study that users of the Virtual Bone Laboratory would need an infrastructure to run simulations and analyze large scale trial data. Hence, Grid services are included in the CORE VRE since such services can provide secure and managed access to distributed computational power.

More details regarding the study may be found in Stenning, Grange, Sim, Wang, Gilbert, and Wills (2005), which describes the user requirements for the CORE VRE in accordance with the study's objectives outlined earlier. Recommendations are also included in the report to guide further development activities in the CORE project.

4. Architecture and Implemented System

The CORE system allows both the laboratory and clinical based users to create technical material (including non research material for education), analyse data (from their own trials or hospital data), investigate hypotheses (from their own work or as meta or thematic reviews), discuss the finding from their or others work, and prepare and submit articles for review. In addition, as some of the analysis from the Virtual Bone Laboratory analysis is computationally intensive, the system also connects to the Grid. The Grid provides access to powerful analysis and modelling tools which are made available to staff and students.

A total of 15 processes were identified from the user requirements and these were factored into four main service areas which were potentially reusable in other VREs. These reusable services are:

Analysis. A Web service to perform analyses on datasets using statistical methods. The processes

- Analyzing data / generating new statistics
- Viewing analyzed statistical results

Data Set Manager. A Web service to handle trial related data: The processes

- Generating trial schemas
- Creating trials
- Viewing / updating trials
- Adding data to experiment logbook

- Viewing data in experiment logbook
- Searching for cases (trial data associated with specific schemas)
- Viewing cases

EPrints A Web service to help users submit and disseminate articles for reviewing between researchers

- Creating preprints
- Viewing preprints
- Updating preprints

Forum. A Web service to support discussions between researchers

- Creating discussions
- Viewing all discussions
- Viewing individual discussions

These services enable users to perform tasks such as formalizing trial protocol, storing and analyzing data, submitting and reviewing articles, and discussing research findings in a forum. A Grid service named *Grid Simulation* is also included in the CORE VRE infrastructure. It provides users with functionalities such as job submission, file transfer, and credential management in running their simulations.

Within the CORE architecture a portal is used to facilitate the sharing of research resources. The main purpose of a portal is to act as a presentation layer which aggregates, integrates, personalizes and presents information, transactions, and applications to the user according to their role and preferences.

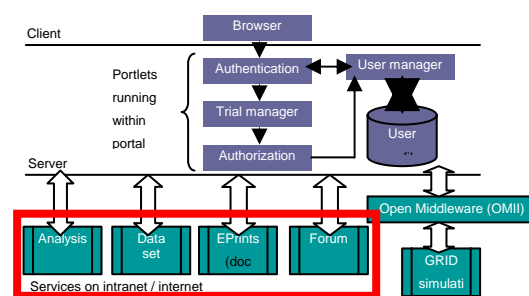
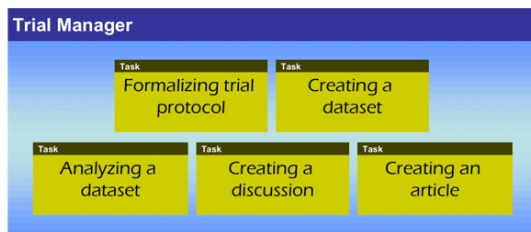


Figure 1 The CORE Architecture

The portal in the CORE architecture employs the OMII middleware so that end users of the VRE can access Grid resources and applications in a trusted and secure environment (Chapman, *et al*, 2005) (Allan *et al* 2004). OMII is a collection of tested, documented and integrated software components that provides a standard platform for integrating e-Science

middleware. It also acts as a simple, secure Web service-based Grid infrastructure for e-scientists.

Figure 2: Overview of the Trial Manager



Within the CORE architecture illustrated in

Figure 1, the *Trial Manager* plays a vital role. It provides the mechanism to configure the VRE in accordance to a particular e-research community's requirements, through the formal specification of e-experiment procedures and workflows relevant to the community. *Trial Manager* performs a set of tasks as depicted in

Figure 2 by using the proposed Web/Grid services in the CORE architecture. To initiate a new trial, a user must first select the experimental protocol from the available experiment schemas defined in the CORE VRE.

An experiment schema lays out the experimental procedures, including recording an experiment description, a statement of purpose, and an outcome hypothesis. The user can then employ the *Dataset Manager* service to create a new dataset for the trial based on a specific data schema. Data schemas describe the exact nature of the experimental data, including specification of variable names, types, and possible values.

To conduct analyses on trial data, users can choose statistical methods offered by the *Analysis* service. Using the experiment schemas and metadata from the *Analysis* service, it is possible to generate a data entry interface for each statistical method.

Having obtained significant results or cited important publications relevant to the trial, users can create a discussion in the VRE. Such a facility is supported by the *Forum* service.

In addition to the tasks described above, the *Trial Manager* also provides facilities for users to create preprints. Publication schemas are used to specify the required format for submitting trial results to relevant journals/conferences. The schemas are used to describe mappings to the trial protocol, for example, specifying that the experiment hypothesis should appear in the *Experimental Methods* section of the preprint. Upon completing the paper, the user can submit it to the VRE by using the *EPrints* service, where the paper becomes available to the members of the community.

5. Related Work

The image many people have of the research process is that of a lone scientist or a small team working in a basement laboratory. A similar picture appears for the use of libraries, where researchers search for vital missing items of information. Levy and Marshall (1995) have examined the early underlying assumptions which have affected digital library development. The images described above was challenged in their article, especially that digital archives/libraries are used by individuals working alone. They argued that the work carried out by researchers and librarians (in providing the service) is one of collaboration, and that digital systems should support formal and informal collaboration and communications.

Similarly, Marchionini suggested that "*digital libraries will allow learners of all types to share resources, time and energy, and experience to their mutual benefit*" (Marchionini *et al* 1995). In their proposed future of digital libraries, sharing resources becomes an important factor in supporting teaching and research activities (e.g. through sharing of scientific data). Many e-Science projects have collected a vast amount of data; nevertheless, if the next generation e-scientists are to go beyond the present position, they will need to have access to the raw data in their research and training. These early visions are slowly being realized, for example McGrath, have developed a system to locate, navigate, and retrieve astronomical data (McGrath, *et*

a/ 1999). However, there is still a need for librarians/archivists and users to work together in producing appropriate tools for handling, manipulating and analyzing these large data sets (Neveill-Manning, 2001).

Marchionini and Maurer (1995) also pointed out that digital libraries should offer greater opportunities for users to deposit information. There are projects beginning to do this, for instance the *Digital Library for Earth Science Education* project provides facilities for students to create and explore geospatial materials and Earth data sets. Duke, Day, Heery, Carr, & Coles (2005) also developed a prototype system for the storage and sharing (via a Web interface) of data collected from crystallography experiments. They envisaged that the need to share the raw scientific data is becoming increasingly essential as e-Scientists often require access to the raw data, and not just the summarized results in published articles, in order to verify the original experiments or to build upon the results.

Shadbolt, Lewis, Dasmahapatra, Dupplaw, Hu, & Lewis (2004) proposed a semantic Web application for the medical domain, the *Triple Assessment Procedure*, for breast screening. Their paper describes the application developed and the knowledge technologies used. However, what is of particular interest is their MIAKT *Task Invocation Framework*. This is a distributed architecture which uses Web services to provide results of distinct and dissimilar processes to a client application. The framework uses a layered model, which consists of a client application, a J2EE Server, and Web services. Although similar to the MIAKT architecture, a slightly different approach is taken in the design of the CORE architecture. A portal is employed to carrying out task allocation, authorization and authentication in the CORE architecture. By using this approach, users can access the VRE by using an Internet browser without the need to install any client applications. There are also developments of e-Science projects in the medical field, such as PsyGrid¹ and IBVRE², sharing the ideology of the CORE

project, which is to build a VRE that combines clinical, educational and research activities under one working environment.

Lyon (2002) sees the digital library, in the context of an information grid, as consisting of a collection of resources for learning and teaching, data repositories for research purposes, or as archives of diverse cultural heritage materials. In her proposed scenario, researchers would undertake experiments, deposit raw data, and produce pre-prints using Web services. On the other hand, Goble and De Roure (2002) presented a case for using the semantic grid to support e-science and particularly virtual research. They state the '*ultimate purpose of the grid is to facilitate knowledge discovery*'. They also provide some scenarios on how Grid and Semantic Web Services can be used to support the sharing of scientific knowledge. Taking the visions and views described above, the CORE project proposed an architecture that is loosely-coupled and designed to meet user requirements.

6. Discussion

The original requirement from JISC was for '*lightweight, proof-of-concept VRE demonstrators appropriate to the needs and skills of specific communities*.'

.....*Specific attention should be given to the skills required to ensure effective exploitation within the target community and any associated training requirements.*'

We believe that we have provided this to the target community of orthopaedic surgeons.

The CORE VRE developed a Web-based portal environment for supporting a critical subset of the e-Science cycle. Also the project was to investigate the use of Grid services using the OMII middleware. We believe that we have provided this to the target community of orthopaedic surgeons. The developed demonstrator enables surgeons to collaboratively:

- create technical material (non research material for education),
- analyse data (from their own trials or data entered from journals),
- investigate hypotheses (from their own work or as meta or thematic reviews),
- discuss the finding from their or others work,

¹ PsyGrid, University of Manchester.
<http://www.psygrid.org/default.aspx>

² Integrative Biology Virtual Research Environment (IBVRE), Oxford University.
<http://www.vre.ox.ac.uk/ibvre/>

- prepare and submit articles for review, using semantic Web services.

The main approach taken in term of the design of the VRE was that of participatory design. This was chosen because of the requirement of the contract for the surgeons to participate through out the life of the project. Participatory design ensured that this main stakeholder group would be involved and engaged through out the development lifecycle.

While developing the demonstrator using the participatory design approach ensured end user involvement, it also caused some tension. The difficulty was that the users wanted the system quickly and were mainly concerned with developing the user interface. So while the use of participatory and co-design approaches are effective in ensuring participation, they do require particularly careful management to ensure the hidden (from the point of view of the user) technical work is also completed (i.e. the infrastructure).

The CORE project developed its own portal, but if we were to start again today we would use one of the many portal frameworks. The quality of the open source portal frameworks has increased over the last two years and with the new portlet standard (JSR286) it may not be necessary for a project to develop its own portal framework. However, it is not as simple to integrate Web services in to a portal framework, as the current CORE portal. The portal framework uses a Model View Control (MVC) approach, and as a result when integrating the Web services code needs to be written to intercept and direct the call to the Web service.

This was a demonstrator project and there is a need to enhance the existing tools, for example the collaborative authoring, especially in light of the recent developments in semantic wikis. With the increase in open source wikis and semantic wikis, embedding these into collaborative authoring tools may be more effective than a traditional Web based system.

The CORE VRE project developed a demonstrator that was trialled in three local hospitals (Portsmouth, Southampton and Salisbury). Technically it was a success, in that all the systems worked as specified and integrated seamlessly.

In addition it met all the users' requirements. There was however one small disappointment; that was the speed with which inexperienced staff could enter data into the system relating to surgical operations. The user evaluation showed that although experienced staff and experienced computer users could use the system effectively, it did require the new user to take time to learn it. This is a fundamental problem with any computerised system, and still a problem in interface design, and it involved the scanning of large multidimensional lists. Current practice is to have lists of the surgical codes for operations entered into a large table. An inexperienced surgeon filing in the paperwork can scan the list for an appropriate code. So these users took longer to use the system than the paper based system.

The demonstrator was designed using an SOA approach, and as such did complete the aims of the project: to demonstrate that a VRE can be build for managing the complexity of multi-centred surgical trial, using Web/Grid services.

7. Conclusions and Future Work

This paper has focused on a specific architecture, the CORE VRE, which supports a critical subset of the e-Science cycle: the collation and analysis of experimental results, the organisation of internal project discussions, and the production of appropriate articles depending upon the requirements of conferences and journals. The CORE project aimed to provide a VRE which is generic and applicable across many e-science domains.

The current system is deployed in the Wessex regional centre for orthopaedics in the UK, and the Orthopaedics department of Portsmouth Hospital. The Grid services are currently being installed in the musculoskeletal laboratory based at Southampton General hospital. All the systems have completed standard system testing, and formal user evaluations have been undertaken. Currently a number of large scale multi-centred trials are being put through the system.

The primary goal of this initiative is to produce an evolving, sustainable, and open standards based service-oriented technical framework to support the education and

research communities. The CORE VRE is a demonstration of how this may be achieved with one particular community.

8. References

- Allan, R., Awre, C., Baker, M. and Fish, A. (2004) Portals and Portlets 2003. Technical report UKeS-2004-06, National e-Science Centre, September 2004.
- Carr, L., Miles-Board, T., Wills, G., Power, G., Bailey, C., Hall, W. and Grange, S. (2004) Extending the Role of Digital Library: Computer Support for Creating Articles. In Proceedings of The Fifteenth ACM Conference on Hypertext and Hypermedi , University of California, Santa Cruz, USA.
- Chapman, S., Dunlop, A., Henderson, P. and Newhouse, S. (2005) OMII Grid Security Technology Overview. Open Middleware Infrastructure Institute, University of Southampton, 2005.
- Goble, C. and De Roure, D. (2002) The Grid: An Application of the Semantic Web. ACM SIGMOD Record, 31(4):65-70, 2002
- Dijkman R., and Dumas M., (2004) "Service-oriented Design: A Multi-viewpoint Approach," International Journal of Cooperative Information Systems, vol. 13, 2004.
- Duke, M., Day, M., Heery, R., Carr, L. A. and Coles, S. J. (2005) Enhancing Access to Research Data: The Challenge of Crystallography. In Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM Press, NY, pages 46-55, 2005.
- Levy, D. M. and Marshall, C. C. (1995) Going Digital: A Look at Assumptions Underlying Digital Libraries. Communications of the ACM, 38(4):77-84, 1995
- Lyon, L.(2002) Emerging Information Architecture for Distributed Digital Libraries. In Proceedings of China Digital Library Conference 2002, 2002.
- Makola D (2006) Evaluation of the South African CPD Information System: using Activity Theory Analysis.
- Marchionni, G. and Maurer, H. (1995) The Roles of Digital Libraries in Teaching and Learning. Communications of the ACM, 38(4):80-81, 1995.
- Martin, J., Arsanjani, A., Tarr, P., and Hailpern, B. (2003) "Web Services: Promises and Compromises," Queue vol. 1, pp. 48-58, 2003
- McGrath, R. E., Futrelle, J. Plante, R. and Guillaume, D. (1999) Digital Library Technology for Locating and Accessing Scientific Data. In Proceedings of the fourth ACM Conference on Digital Libraries, pages 188-194, 1999.
- Neveill-Manning, C. (2001) The Biological Digital Library. Communications of the ACM, 44(5):41-42, 2001
- Quartel D. A. C., Dijkman R. M., and van Sinderen M. J., (2004) "Methodological Support for Service-oriented Design with ISDL," in proceedings 2nd ACM International Conference on Service Oriented Computing (ICSOC), New York City, NY, USA, 2004.
- Shadbolt, N., Lewis, P., Dasmahapatra, S., Dupplaw, D., Hu, B. and Lewis, H.(2004) MIAKT: Combining Grid and Web Services for Collaborative Medical Decision Making. In Proceedings of the UK e-Science All Hands Meeting 2004, Nottingham, UK. <http://www.allhands.org.uk/2004/proceedings/papers/258.pdf>
- Stenning, M., Grange, S., Sim, Y. W., Wang, C., Gilbert, L. and Wills, G. B. (2005) CORE User Requirement Study. Technical report ECSTR-IAM05-007, University of Southampton, 2005. <http://eprints.ecs.soton.ac.uk/11229/>
- Wilson, S., Blinco, K., & Rehak, D (2004) Service-Oriented Frameworks: Modelling the infrastructure for the next generation of e-Learning Systems. Joint Information Systems Committee (JISC). Bristol, UK 2004.