Advance & attack: technology demonstrator to the rescue … and beyond

Dr Martie van Deventer (CSIR)
&
Dr Heila Pienaar (UP)

10th Internet Librarian International Conference, London, 16 October 2008
Roadmap

- Context and our research
- Demonstrate the demonstrator
- Convincing Executive
- Way forward
Context

- South African Research Information Services (SARIS) project – 2004
  - Identified virtual research environments (VRE’s) as an important component of current global research – no national initiative was forthcoming

- Utilising the SERA relationship between CSIR & UP
  - SERA – Southern Education Research Alliance
  - Our aim was to establish a conceptual framework
  - Needed research area with much data generation
  - African Centre for Gene Technologies’ South African Malaria Initiative (ACGT SAMI) was identified by Executives

- Malaria VRE research project
  - ACGT management agreed to participate
  - Completed ‘a day in the life’ & research tools semi-structured interviews with a variety of malaria researchers in the SAMI network to establish their readiness to move to an integrated VRE
Malaria: Endemic / Epidemic Risk Areas

Source: http://www.mara.org.za/mapsinfo.htm
Definition - VRE

• The VRE concept helps to broaden the popular definition of e-science from grid-based distributed computing for scientists with huge amounts of data to the development of online tools, content, and middleware within a coherent framework for all disciplines and all types of research (Fraser, 2005).

• The specific aim of a VRE is to help researchers manage the increasingly complex range of tasks involved in carrying out research. Therefore a VRE provides a framework of resources to support the underlying processes of research on both small and large scales, particularly for those disciplines which are not well catered for by current infrastructure (JISC, 2006).
Definition - Demonstrator

• A technology demonstrator is defined as an output of an R&D project that is in a state of technology readiness to enhance capability, manufacture a product or deliver a service in a previously unknown manner.

• Examples of technology demonstrators include a novel machine to extrude polymers, a novel process to manufacture a chemical product, a novel software application and a new process (key solution) to improve decision support or the operation of major organisations and government. Novel research tools that improve R&D capability or productivity are also included in the definition (CSIR).
“A day in my life”

<table>
<thead>
<tr>
<th>Institution</th>
<th>#Interviewed</th>
<th>Male</th>
<th>Female</th>
<th>Management</th>
<th>Research</th>
<th>&lt;30</th>
<th>30-40</th>
<th>&gt;40</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSIR</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>UP</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other SAMI</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>20</strong></td>
<td><strong>13</strong></td>
<td><strong>7</strong></td>
<td><strong>8</strong></td>
<td><strong>12</strong></td>
<td></td>
<td><strong>1</strong></td>
<td><strong>7</strong></td>
</tr>
</tbody>
</table>

Age group
Researchers

- Start the day with e-mail
- Largest chunk of day spend in wet lab
- Articles are written by teams
- Management of research data: majority of files only traceable via the lab book
- Electronic lab books?
- Much time spending on report writing ….
Managers

• Days are planned around scheduled meetings
• Face-to-face their preferred mode of communication
• Spend “alone time” looking at research agendas, trends and opportunities
• Problems with malaria could only be resolved by multi-disciplinary teams ....
Identification of research area

Dissemination of findings (artifacts)

Real time communication

Training / mentoring etc

Scientific workflow

Literature review & indexing

Identification of collaborators

Proposal writing

Identification of funding sources

Project management

Sub-routine to test ideas

IP Management was not identified but needs to be taken into consideration
Identification of research area:
- Personal networks; face-to-face; literature; government documents

Identification of funding sources:
- Personal networks; funding agencies; institutional resources e.g. SAMI

Identification of collaborators:
- Personal networks; EU portal; literature; search engines; ACGT expert list

Proposal writing:
- MS Word / Open Office; templates; generic proposal

Literature review & indexing:
- Preferred databases: PubMed, Science Direct, Scopus; Retrieval: Google Scholar, Browser favourites; Filing: manual; database

Dissemination & artifacts:
- E-mail; face-to-face; phone; webex; wiki; website; meetings

Real time communication:
- Face-to-face; hands on; UP: e-Learning for students

Training / mentoring etc:
- Sophisticated instruments with own software – write data to servers; Free analysis software; Paper lab book; Referencing system between lab book and instruments; Ad hoc management of data (curation)

Scientific workflow:
- CSIR: formal pm with tools & staff; UP: informal

Project management:
- CSIR: formal pm with tools & staff; UP: informal

Current practices & tools
SAMI wish list

Identification of research area

Grouping of info in one place

Identification of funding sources

List of funders easily accessible e.g. web site

List of search engines; Internal shared database of indexed articles; Person to assist in retrieval of relevant literature

Literature review & indexing

Identification of collaborators

Proposal writing

Document management system

List of researchers & topics

Project management

Proper pm system; MS Project

Scientific workflow

Even more sophisticated instruments; Electronic lab book; Systems biology software; Experiment repository; Labs with in silico screening+; Bio-information specialist

Dissemination & artifacts

Real time communication

Training / mentoring etc

e-Learning system for researchers

Skype; collaborative e-Lab books; Smart board; video conf; project portal

SAMIS wish list
Consolidated SAMI VRE components

Repositories: research results; experiments; literature & documents

Web/wiki/blog: search engines, databases; researchers & topics; funders, portals, communication, projects

Real time communication

Skype, smart board, video conferences

Identification of research area

Identification of collaborators

Training / mentoring etc

Proposal writing

Identification of funding sources

E-learning system for researchers

Scientific workflow

Document management system

Internal shared database of indexed articles

Dissemination & artifacts

Real time communication

Repositories: research results; experiments; literature & documents

Mathematical modelling tools; numerical algorithm tools; simulation software; in silico experiments

Integrated data management system

Sophisticated instruments that generate digital information and data

Access to research networks & super computers; access to labs with in silico screening +

(Free) Data analysis software

Project management system
None of the researchers indicated that they had access to:
- A web / wiki / blog to use for lists of search engines, databases, researchers, funders, portals, projects, software, instruments
- Repositories for research results (articles, data etc), experiments and documents
- An integrated data management / curation system
- Collaborative electronic lab book system
- An e-Learning system for researchers, e.g. to transfer knowledge about new methodologies

Only some of the researchers are making use of:
- An internally shared database of indexed articles. Individual databases (paper / electronic) are quite popular.
- A document management system (CSIR)
- A project management system (CSIR)
- Access to research networks, super computers and labs with in silico screening+
- In silico experiment software
- Electronic communication tools (Skype, Smart board, Video conferencing etc)

And all of the researchers are making use of:
- Sophisticated instruments that generate digital information and data
- Servers with data files
- Mathematical modeling tools
- Numerical algorithm tools
- Simulation software
- Data analysis software (mostly freeware)
- Generic software e.g. MS and Open Office
The demonstrator
Building the VRE demonstrator

• This was seen as a mock-up model – not a pilot nor the start of a working VRE
• Made use of third year UP Information Science students (practical work) and CSIR Interns (who were already populating our institutional repository using DSpace)
• Development happened over a two week period
Web 2.0 tools used

- Proposal writing
- Identification of research area
- Identification of funding sources
- Literature review & indexing
- Scientific workflow
- Real time communication
- Commercial Resources via Library
- Identification of research area
- RSS & Alerts
- IP Management was not identified but needs to be taken into consideration
- Training / mentoring etc
- Media Wiki
- MSM/ Google IM/ GMail
- Combination of Blogger & DSpace
- Dissemination & artifacts
- Proposal writing
- Project management
- WebCoLab
- Portal interface
- Xoops

WebCoLab

Identification of collaborators/ shared resources/ Experts

Commercial Resources via Library

RSS & Alerts

MSM/ Google IM/ GMail

Media Wiki

Combination of Blogger & DSpace

Dissemination & artifacts

Scientific workflow

IP Management was not identified but needs to be taken into consideration
Interface

Access and authentication

Library resources

Internet resources

Alerts and RSS feeds – may help with info overload
Authorisation through application

Commercial Resources via Library

Personal space

Blogger

Wiki

DSpace

WebCoLab

Shared space

MSM/Google Groups

Welcome

The South African Malaria Initiative (SAM) is a partnership established to facilitate the integration of malaria research capacity development in South Africa and the rest of Africa. The aim is to stimulate the use of modern molecular research tools to improve malaria prevention and control.

The South African Malaria Initiative will facilitate an integrated programme of malaria research an Africa and eventually in the rest of Africa to improve malaria prevention and control. Modern malaria research outputs will include the identification and validation of drug and insecticidal targets, development of drug and insecticidal candidates, improved diagnostics, and new tools for gathering epidemiological information.

Latest publications on malaria

Click here to run an automated Google Scholar search

(Click here to go to malaria Journal)
The repository

'Artefact type' collections
Lab book

Linda's CSIR Biotechnology lab-book

This is Blog for Professor Linda Mtswisha. She volunteered to lend her lab-book (for Malaria) in order for us to demonstrate virtual research environment, not everything is accurate. Roller was not working so we subscribed to the www.blogger.com. The outcome of the project is fruitful.

Friday, July 13, 2007

Cartoon

"You should check your e-mails more often. I fired you over three weeks ago."
Wiki

Wiki stored the information the group regards as their reference material.

Instructions for use

In addition, each VRE user has their own Profile Page, where they can give as much or as little information about themselves and their collections as they like.

There is a page listing tools for working with data provided by VRE users.

To get started here, take a look around and start your own page or contribute to an existing one. If you would prefer to practice editing before you work on a real page, you can try it out in the sandbox. When you edit a page, you will see more information on how to format text, and if you need any more information, see the help pages. Any problems, ask other users on these pages or email us at mvandev@csir.co.za or hella.pienna@up.ac.za
Executive discussion

- **Easy** (Web 2.0)
  - Web / Wiki / Blog
  - Repositories / database of indexed articles/ data sets
  - Communication devices
  - Project management

- **Important**
  - Integrated data management system
  - Collaborative electronic lab book
  - Combining these two applications
Value of the demonstrator

- Visual – the discussion moved from theory to practice
- VRE appeared implementable & useable
- Efficiency gains were easy to understand – it is quite clear that researchers would be able to work smarter
- Effectiveness in terms of new ways of research has become the point of discussion and not the applications (able to move from ICT to content)
- Ability to curate the organisations’ digital content has become more feasible (even do-able!)
Recent international VRE-type initiatives

- Integrative Biology VRE
- **HUBzero** generic science gateway platform (will become open source in 2009)
- OPENWETWARE initiative
- **British Library** Research Information Centre beta software
- ResearchGATE science network (science 2) platform
- Taverna workbench software for the biosciences, and
- Project 35 open source software for data entry and modelling
- VRE for the **Humanities**
WHAT IS HUBZERO?

HUBzero allows you to create dynamic web sites that connect a community in scientific research and educational activities. HUBzero sites combine powerful Web 2.0 concepts with a middleware that provides instant access to interactive simulation tools. These tools are not just Java applets, but real research codes that can access TeraGrid, the Open Science Grid, and other national Grid computing resources for extra cycles.

HUBzero was created by researchers at Purdue University in conjunction with the NSF-sponsored Network for Computational Nanotechnology. The technology was originally developed to support nanoHUB.org, a national resource for nanotechnology simulation. It has since been extended to create science gateways for other scientific domains.

View a demo of nanoHUB.

Learn more about the underlying infrastructure.

HUB Watch

Here is a list of hubs being built with the HUBzero software:

- cceHUB.org
case centered engineering
Sparse random networks of carbon nanotubes with thousands of tubes and random orientations. Compute the conductivity of these networks and simulate their performance in transistors of various geometries.
VRE for the Humanities

• Building a virtual research environment for the Humanities (BVREH). An initial survey carried out by the BVREH team between June 2005 and September 2006 defined the range of services that a Virtual Environment should offer - from information about researchers and their interests and about conferences, lectures and seminars, to integrated communication and collaboration tools to support advanced research. The "Digital Pen and Paper Technologies" project is showing a lot of promise (http://bvreh.humanities.ox.ac.uk/).

• The VRE for the Study of Documents and Manuscripts project naturally follows from the outcomes of BVREH project. In this project a broad-based understanding of user-driven needs has been established, and it was shown how tools and resources for studying texts and document might be implemented in a service-based environment and some annotation and mark-up tools have also been tested. The project team will now proceed to construct an integrated environment in which the data (documents), tools and scholarly instrumenta will be available to the scholar as a complete and coherent resource (http://www.jisc.ac.uk/media/documents/programmes/vre2/vre18sdmvreprojectplan.pdf ; http://www.jisc.ac.uk/whatwedo/programmes/programme_vre/vre_bvrehs.aspx ; http://bvreh.humanities.ox.ac.uk/VRE-SDM ).
Way forward

- Many new developments since we started the conceptualisation project; possibility of open source availability
- ACGT manager has acquired the necessary funding to start building a pilot
- Project will now become a collaborative effort where the researchers determine progress
- We’ll be advising on the available platforms & information services that could serve as basis for the development
- **Possibility to develop a generic VRE for SA**
References

- myExperiment: [http://www.jisc.ac.uk/whatwedo/programmes/vre2/myexperiment.aspx](http://www.jisc.ac.uk/whatwedo/programmes/vre2/myexperiment.aspx) [last accessed 09 July 2008]
- OpenWetWare ([www.openwetware.org](http://www.openwetware.org)) [last accessed 09 July 2008]
Questions / Comments?

Dr Martie van Deventer  
mvandeve@csir.co.za  
012 841 3278

Dr Heila Pienaar  
Heila.Pienaar@ais.up.ac.za  
012 420 2020