Supporting the e-Research lifecycle from acquisition through to annotation: the DART/ARCHER experience

Dr Andrew Treloar
DART Project Architect
ARCHER Director and Chief Architect
ARROW Technical Architect

Presentation to eResearch Australasia 2007
First some context

Grid

E-Scientists

Entire E-Science Cycle
Encompassing experimentation, analysis, publication, research, learning

Virtual Learning Environment

Institutional Archive

Local Web

Publisher Holdings

Digital Library

E-Experimentation

Technical Reports

Preprints & Metadata

Local Web

Certified Experimental Results & Analyses

Data, Metadata & Ontologies

Preprints & Conference Papers

Peer-Reviewed Journal & Conference Papers

Source: Adapted from Liz Lyons, eBank UK Presentation
DART

- DART is/was a proof-of-concept project funded by the Department of Education, Science and Training (DEST) to support collaborative research in Australia.
- The funding has been provided through the Systemic Infrastructure Initiative as part of the Commonwealth Government's Backing Australia's Ability - An Innovation Action Plan for the Future.
- DART stands for:
  - Dataset
  - Acquisition/Accessibility/Annotation
  - e-Research
  - Technologies
DART logistics

- 3 partners:
  - Monash University (lead) in Melbourne
  - University of Queensland in Brisbane
  - James Cook University in Townsville
- 5 technical areas of focus within the DART work packages (WPs)
- 18 month project, finishing in June 2007
- 28 Separate DART work packages
- 40+ project team members!!
DART chief investigators

Andrew Treloar
Asad Khan
David Abramson
Ann Monotti

Jane Hunter
Xiaofang Zhou

Ian Atkinson
What did DART try to achieve?

- To develop software tools to handle the data and information management requirements of the complete research lifecycle
- To collect and manage large datasets, associated with instruments, such as sensor networks, X-ray diffractometers, etc.
- To support collaborative research and annotation needs
- To deal with intellectual property, privacy and security issues
- To create customised portals for research demonstrators
- To handle research publication, discovery and access

or to put it another way......
DART work packages

- The work packages cover five broad technical areas:
  - Data Collection and Monitoring
  - Storage and Interoperability
  - Content and Rights
  - Annotation and Assessment
  - Discovery and Access

- Details [on the website](http://dart.edu.au), including all project outputs:
  - reports, source code, documentation

- Or come and see the stand in the exhibition area!
DART Achievements

- Strong progress in data capture and instrument integration
- Investigated storage and replication of very large datasets across diverse networks
- Placed Information Management staff into research teams, addressing their data and information management requirements
- Developed annotation software for 3-D models, video and audio
- Reviewed IP and privacy around datasets and e-research practices
- Investigated Creative Commons & Science Commons licensing
- Worked to utilise Shibboleth, PKI and Grid security standards
- Developed search tools, metadata schema registry, Plone tools, etc
- Production deployment of X-Ray Crystallography tools
DART Demo

- NOTE: This is only a partial demo of the range of work performed in DART - for more see the website or visit the stands in the exhibition area

DART Lessons

- Importance of demonstrators
  - not envisaged in original bid
  - demonstrating end-end benefits
- Integration challenges
  - didn’t originally employ someone to focus on this
  - underestimated amount of complexity and effort required
- Collaboration and co-ordination
  - Tendency to focus on the specific task
  - Problem of distributed activity
  - Lots of travel around
  - Over-busy chief investigators
- Rate of technology change
DART Conclusions

- Value of end-end approach didn’t evidence itself as much as expected
- Different groups get excited about particular components
- Researchers are very time-poor and need immediate value to stay involved
- Need agile development techniques to deal with loose or missing requirements
- Some of the changes we envisaged in DART are now happening in scholarly communication practices and national initiatives
  - RQF accessibility component
  - NCRIS Platforms for Collaboration (ICI and ANDS)
From DART to ARCHER

- Australian ResearCH Enabling enviRonment

ARCHER is a new DEST funded project for 2007 that will take the proof-of-concept outcomes of DART, turn them into production-ready ARCHER software tools, and package them, along with other open-source components, for deployment.

- These tools are being developed as modular middleware components, customised to suit the needs of a number of designated National Collaborative Research Infrastructure Scheme (NCRIS) priority research capabilities.

- Now engaging with NCRIS capabilities and undertaking active development.
ARCHER contraction of scope

Grid

E-Scientists

Entire E-Science Cycle
Encompassing experimentation, analysis, publication, research, learning

Virtual Learning Environment

Institutional Archive

Local Web

Publisher Holdings

Digital Library

Peer-Reviewed Journal & Conference Papers

Reprints

Preprints & Metadata

Technical Reports

E-Experimentation

E-Scientists

Graduate Students

Undergraduate Students

Data, Metadata & Ontologies

Certified Experimental Results & Analyses

Source: Adapted from Liz Lyons, eBank UK Presentation
Acknowledgements

- Without the hard work of all these people (and more!), DART could not have happened!
Questions?

- Web:
  - dart.edu.au
  - archer.edu.au

- Email:
  - DART/ARCHER: andrew.treloar@its.monash.edu.au