The Australian Ocean Data Network:
Building the e-Research Marine Science Information Infrastructure

Craig Johnson
University of Tasmania

Our Vision

A single distributed, federated network for Australian marine science data and information.

To:
• Promote the discovery, transfer and implementation of knowledge about marine environments by facilitating the management and exchange of, and ready access to, marine data and information.
• Provide a key enabling technology for e-research.
Building the AODN

- AODCF as 'core component'
- BlueNet project
- eMII - project of IMOS (Integrated Marine Observing System: NCRIS)

The AODCF

Australian Ocean Data Centre Joint Facility

A partnership among Commonwealth Agencies:
- Australian Antarctic Division (AAD)
- Australian Institute of Marine Science (AIMS)
- Bureau of Meteorology
- CSIRO Marine and Atmospheric Research
- Geoscience Australia
- Royal Australian Navy

- with support from the National Oceans Office

To manage ocean data to meet national and international obligations
BlueNet

A DEST-funded project with three main roles:

1. Join universities to the AODN
   - data facilitators
   - BlueNet Champions

2. Help agencies prepare to host 'external' data
   - building the virtual repository

3. Develop technologies for the AODN
   - Metadata Entry and Search Tool (MEST)
   - support for Oceans Portal / Marine Catalogue

BlueNet - partners

- AODCJF Agencies, TPAC
  and
- James Cook University
- University of Sydney
- Australian National University
- University of Melbourne
- University of Tasmania
- University of Western Australia

- (University of Queensland)
- (Deakin University)
The Network (AODN)

The data management component of IMOS

Australian Ocean Data Network
AODN Technologies
Portal & Catalogue

Portal

• Visualises / integrates disparate datasets
• GIS-like functionality (map layering, legend, zoom, etc)
• GUI to search for and call services listed in Catalogue
• 4 core types of searches (Species, GCMD topics, Spatial, Freetext)
• Allows user to build complex searches

Catalogue

• Harvested metadata records are imported into the Catalogue
• No GUI; handles calls from remote clients (e.g. Oceans Portal) for searching
• Is a registry for:
  – metadata
  – data
  – for web services (provides links to services that facilitate data access, data manipulation, visualisation and download)
Metadata Entry & Search Tool (MEST)

- Federated searching, records harvesting, and batch import
- Web map services - allowing visualisation of multiple datasets
- Manages permissions and access (user levels for metadata viewing, creation and editing, and data access)
- Context-based and content-based display of "Copyright and Use-limitations"
- Gathers statistics about data-downloads; data owners advised of the users of their data
- ISO19115 compliant (the international geospatial metadata standard) plus other metadata standards
- Offers several vocabularies
- Customisable metadata-entry templates for varying data types
- Data upload and download
- Deployed on a central server and at data-hosting agencies (for local needs)
The Network (AODN)

AODN - Outreach

Provision of tools and services for BlueNet and eMII:
- Comprehensive help documentation – for MEST use, metadata creation, IP, hosting issues, etc
- “In-the-field” tool, and conversion program, for non-networked metadata entry
- User guides, induction guides, web pages, FAQs
- Discussion lists
- Help desk
- Workshops and training: standard and special-needs sessions (for various stakeholder groups)

Services to maximise data exposure/ discoverability
- Publish metadata to the Australian Spatial Data Directory
- Links to other similar information resources and services
Research projects are complex and dispersed, resulting in a need for:

- Continued awareness of the initial and changing commitment of partners to project goals
- Appreciation of the external drivers on the participants
- Regular face-to-face meetings to reinforce participation and commitment, and to communicate plans / actions / tasks
- Careful management of remote staff to maintain focus
- Keeping an eye on the true goals – the system and the infrastructures are means to an end, not an end in themselves

Challenges:
the complexities of an e-research project

Changing Culture - Data Sharing

- Educating scientists about the utility of their science (legacy = data as well as papers)
- Work flows at an institutional level
- Overcoming researcher resistance to depositing data:
  - reservations about sharing data (loss of professional advantage; concerns about misuse by others)
  - concerns about the time needed to input metadata/data
- Need for credit for making data available
- Acknowledgment of data reuse
Challenges: the complexities of an e-research project

Changing Cultures - Institutional Issues

• Interactions with pre-existing programs and practices
• Different understanding and capacity within an organisation (policy at top not related to staff view)
• Priorities in resource allocation in an environment of change management
• Changing roles of hosting organisations

Standards !@#$??@!

• Data, metadata, QA / QC, timeliness, archival practice etc.
• An ‘end-to-end’ problem in the data flow
• Getting agreement between disparate bodies, e.g. in:
  • developing metadata standards, vocabularies
  • interpretation and application of standards
• Integration of standard practices with pre-existing programs
  • standards developed for the AODN have to mesh with those in pre-existing, data management-mature programs
Building and managing expectations

• Consistent views of and values about broader goals and deliverables
  • funding body’s view
  • development team’s view
  • scientist’s view

• Developing a common vision among planning, development and implementation teams

• Faith needs to be developed in pioneering projects harnessing new technologies

• Changing cultures is a non-trivial task that takes time

Challenges:
the complexities of an e-research project

Technical & other challenges

• AAA, dual identities, and transitory organisational membership

• Standards development (and adoption)

• Coping with different software [proprietary, legacy, etc] for data storage, management etc

• Developing web services: web-map services, web features etc

• Versioning, duplication and ‘point-of-truth’:
  • data (raw, QC’d, integrated, real time vs. delayed, delayed error-correction)
  • metadata

• IP