

SCAR Scientific Data Management

Data and Information Strategy

Scientific Committee on Antarctic Research



Background

- SCAR Executive direction to JCADM - develop a SCAR Data and Information Strategy.
- Begun in 2006 - post Hobart SCAR Conf.
- Canvassed SCAR science programs in 2007 for input.
- Draft Strategy developed in 2007.
- JCADM Workshop in Rome 2007.
- Finessed in 2007/2008 and provided as a draft to EXCOM in 2008 prior to St Petersburg Meeting. Circulated to delegates for comment.
- EXCOM established *Ad-hoc* Action Group post St Petes to review and finalise Strategy. Broad circulation and call for comments followed.
- Action Group to provide final Strategy document to EXCOM in August 2009.
- Post 2009 development of an implementation plan.



Ad Hoc Committee Members

- SCAR V-P for Science (S Marensi) (Chair)
- Representative of SSG-LS (Kathy Conlan)
- Representative of SSG-PS (Maurizio Candidi)
- Representative of SSG-GS (Terry Wilson)
- Representative of SC-ADM (Kim Finney)
- Representative of SC-ADM (Helen Campbell)
- Representative of SC-AGI (Henk Brolsma)
- External Representative (Lesley Rickards, past Chair of IODE, and Chair of the JCADM review group)



Benefits of a Strategy

- ATCM resolutions urging parties to:
 - Support Antarctic Metadata Directory (AMD)
 - Establish NADCs
 - Develop an Antarctic Data Management System (ADMS)
 - Submit data to all of the above.
- ToRs for SCADM (JCADM) include:
 - Development of the ADMS
 - Definition of the “ADMS” as some “system” based on AMD and NADCs.
- Scientists themselves expressing frustration that its hard to find and get access to data of interest.



Benefits of a Strategy

- Problem is:
 - No guidance on what an NADC “is” or what it should “do”.
 - No power to “obligate” establishment of NADCs.
 - No shared view of what an “Antarctic Data Management System” is.
 - No truly “coordinated” approach to managing SCAR related “data”.
 - In the absence of the above, people doing their own thing. Exception is use of the AMD.



Benefits of a Strategy

- Dictionary Definition of a “system”:
 - “instrumentality that combines interrelated, interacting artifacts designed to work as a **coherent entity**”
- Systems don’t emerge fortuitously you need to design and build them - requires planning and coordination.
- Strategy should address current shortcomings in our approach to data management, particularly development of our “system”.



Key Ingredients Of A Useful Data (Management) System

Archive Facilities

- Requires:
 - Community agreement on designated archiving Centres or Institutions for data retention, geared to provide ongoing access and re-use of data.

(Scientific Data Centre networks already exist which could be leveraged for archiving SCAR science data)



ICSU World Data Center System USA Home Russia Home UK Home

ICSU > WDC Home



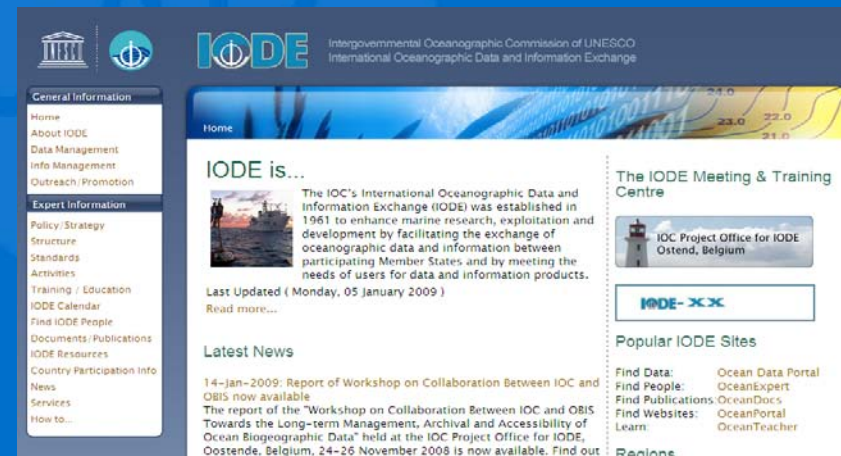
World Data Center System

Data constitute the raw material of scientific understanding. The World Data Center system works to guarantee access to solar, geophysical and related environmental data. It serves the whole scientific community by assembling, scrutinizing, organizing and disseminating data and information.

LATEST NEWS ON THE NEW WORLD DATA SYSTEM

Please note that, as of the end of 2008 following the ICSU General Assembly in Maputo, Mozambique, the Panel on World Data Centers no longer exists and that the World Data Centers will be incorporated into the new ICSU World Data System (WDS) in 2009. A similar procedure will involve members of the Federation of Astronomical and Geophysical data-analysis Services (FAGS) grouping of ICSU. In addition, it

- About the WDC System
- List of Current WDCs
- Contact a WDC
- Search for WDC Data
- WDC System Guide



IODE Intergovernmental Oceanographic Commission of UNESCO International Oceanographic Data and Information Exchange

Home

IODE is...

The IOC's International Oceanographic Data and Information Exchange (IODE) was established in 1961 to enhance marine research, exploitation and development by facilitating the exchange of oceanographic data and information between participating Member States, and by meeting the needs of users for data and information products.

Last Updated (Monday, 05 January 2009)
Read more...

Latest News

14-Jan-2009: Report of Workshop on Collaboration Between IOC and OBIS now available
The report of the "Workshop on Collaboration Between IOC and OBIS Towards the Long-term Management, Archival and Accessibility of Ocean Biogeographic Data" held at the IOC Project Office for IODE, Oostende, Belgium, 24-26 November 2009 is now available. Find out

The IODE Meeting & Training Centre

IOC Project Office for IODE
Oostend, Belgium

Popular IODE Sites

Find Data:	Ocean Data Portal
Find People:	OceanExpert
Find Publications:	OceanDocs
Find Websites:	OceanPortal
Learn:	OceanTeacher

Regions



Key Ingredients Of A Useful Data (Management) System

Ability To Discover Data

Requires:

- Systems and people available to manage/publish data and metadata



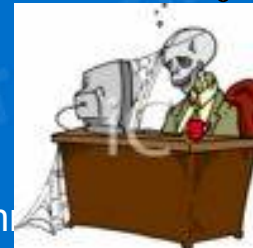
- Investment by scientists to write metadata and systems that make that task easy and efficient



This all seems too hard.
How much information is
enough !

5,000 metadata
entries for climate
and not one has any
real data attached !

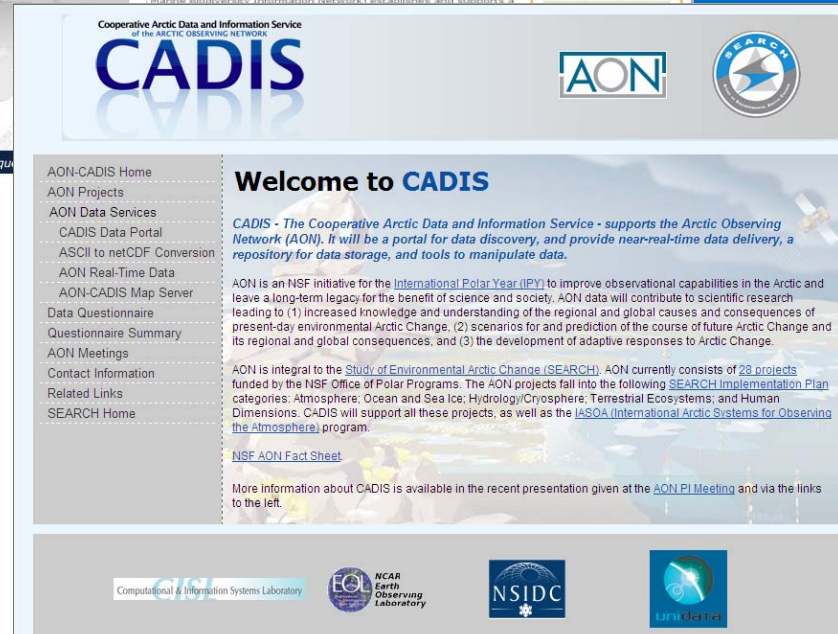
- Ideally metadata and data should be bundled or linked



Key Ingredients Of A Useful Data (Management) System

Ability To Easily Access Data

- Requires:
 - Suitably maintained sites/tools, with intuitive user interfaces.
 - Willing data providers and motivated data users.
- Many existing & newly establishing systems that could be leveraged.



What Is A Tractable Strategy Given Our Context ?



First Strategy draft has been labelled a “Rolls Royce”.

Perhaps what we can afford is closer to a mini minor.

Which makes it all the more important to drive on someone else's super highway



Superhighway Components We Can Harness:

- National Data Centres
- Global Data Centres
- Thematic Data Centres
- Specialised Repositories
- Publishing Networks
 - Protocols
 - Servers
 - Standards
- Virtual Observatories
- Specialised Data Portrayal Software
 - Mapping
 - Plotting, time series

Loosely stitch together what we need.



Strategy

- SCAR Strategy - not a SCADM Strategy.
- Mentions SCADM and SCAGI a fair bit because these are SCAR groups charged with addressing SCAR's data issues.
- But looks outward to formalise partnerships with credible, existing players on the global data scene.
- Strategy is “aspirational” but anticipates that implementation will be incremental.



Current Strategic Suggestions

- Formalise a Data Policy
 - Access to data,
 - Promoted publication networks (data and metadata),
 - Profile functions and obligations of NADC's,
 - Data management planning as part of SCAR project proposal process,
 - Archiving guidelines (principles, facilities)
- Leadership & Funding
 - Explore possibility of dedicated resourcing to drive Strategy.
 - Explore opportunities to fund/resource Strategy implementation (grants, national programs, foundations).
- Culture
 - Explore opportunities for developing or accessing a data citation system.



Current Strategic Suggestions

- Operation of SC-AGI and SC-ADM
 - Rationalise use of face-to-face meetings,
 - Greater cross-fertilisation between groups (possible merger ?)
 - Improved communication and liaison with science to provide services required
- NADC's
 - Explore possibility of establishing regional/thematic data hosting facilities for data publication and data archiving.
 - Small pilot projects could be used to establish/develop services.
 - Explore partnership with other Data Centre/science networks (e.g. WMO, WDS, IODE, IPYDIS, Arctic networks, specialist networks).
 - Foster emerging and less mature NADC's through training and guidance.



Current Strategic Suggestions

- Map-centric data
 - Expand use and application of existing GI products (e.g. Gazetteer, Map Catalogue, Digital Database)
 - Explore gaps in topographic data coverage and means of more collaborative capture and product development, particularly in science hot-spots.
- AMD
 - Better feedback to GCMD on SCAR user needs and greater involvement in strategic planning of future directions for AMD.
 - Pursue strategies that get more data linked to metadata.



Questions and Opinions ?

Thank-you for your
patience and attention.

Your views will be fed back to the
Data Action Group.



Spare slides - if required

Scientific Committee on Antarctic Research



Current Strategic Suggestions

All framed under ToRs for SC-ADM and SC-AGI

SC-ADM

1. To promote long-term preservation and accessibility of data relating to Antarctica and the Southern Ocean in sustainable repositories,
2. To assist in establishing Antarctic data management policies, priorities and best practices,
3. To support the establishment and ongoing work of National Antarctic Data Centres, in accordance with ATCM XXII Resolution 4.1 (1998),
4. To encourage submission of metadata and data to the Antarctic Data Management System,
5. To further improve and populate the AMD and provide guidance to the AMD host,
6. To provide linkages to other relevant data management systems and thereby enhance the ADMS,
7. In partnership with SCAGI to work with SCAR SSGs, COMNAP and the Antarctic Treaty Secretariat to identify and develop fundamental datasets of value to the Antarctic Community.

SC-AGI

1. Provide Antarctic fundamental geographic information products and policies in support of science programs.
2. Integrate and coordinate Antarctic mapping and GIS programs.



Action Group ToRs

- to revise the draft SCAR Data and Information Management Strategy document, taking into account feedback from the wider community, and noting that Delegates at XXX SCAR considered that the present document mixes strategy and implementation, which need to be teased apart, and focuses unduly on JCADM and SC-AGI without sufficient attention to other SCAR data activities;
- to ensure that the strategy positions SCAR in relation to the larger picture developing globally, including *inter alia* taking into consideration coordination of developments on data sharing within ICSU, IOC and WMO;
- to consider the roles SCAR could realistically take on in regard to these activities as an IPY Legacy once the IPY IPO is disbanded;
- to ensure to the extent possible (through consultation) that there is a consensus from other major players in Antarctic and Southern Ocean data and information management (including the user communities) on any recommendations for SCAR actions (these linkages, partnerships and cooperations should be an explicit part of the overall plan);
- to report on progress to EXCOM in Punta Arenas in mid-2009.
- Once a revised strategy is agreed, to prepare a phased implementation plan clearly outlining the timing and extent of any resource implications and, the potential sources for such funding, for presentation to XXXI SCAR (Buenos Aires, 2010)
- to optimize utilization of limited funds and participation the group should work primarily by e-mail, and meet in association with the Cross-SSG Workshop (late 2008 or early 2009).



Why Would We Want A Data Strategy ?

- To improve scientists awareness of available data (particularly for multi-disciplinary studies). Data is expensive to acquire in Antarctic region - re-use can reduce costs - important in hostile economic climate.
- Re-use is feasible when existing data is documented. Little value if quality is unknown.
- Large, heterogeneous data sets now routinely collected via remote methods. Better data handling techniques are required to utilise and integrate these data with other sources. More data being acquired than there are scientists to analyse them. Sharing/collaborative projects make most sense.
- Unequal distribution of resources and skills within SCAR community. More capable players can assist less well resourced with practices and technologies but in turn benefit by getting access to more data.
- Pooling efforts in providing data services in strategic areas provides for efficiencies in resource-limiting times. Many groups within SCAR and external bodies (e.g. COMNAP, ATCM, CEP CCAMLR) share common data-centric problems that can best be solved collaboratively. Politics, trust and individualism are larger barriers to overcome than the technology.



Why Would We Want A Data Strategy ?

- Lots of “bits” of what we need already exists outside of our community, or on the fringes of it, if we are clever we can leverage the work that others have done to our advantage (and be welcomed for doing so).
- Any type of infrastructure needs to be planned and built, rarely do they simply ‘emerge fortuitously’. If we want a better data infrastructure some degree of planning and coordination is necessary.
- New fields of science are appearing in informatics and modelling to take advantage of the growth in data and computing power. This requires handling data in new and more automated ways if we want to get the fruits of what this type of science can deliver.
- Younger scientists are ‘Net savvy’ and expect to share and acquire data on-line.
- Our current approach will not deliver on most people’s expectations. We must agree on what we want to achieve and then how to get there.

