—シンポジウム/会合報告—
Symposium/Meeting Report

Report of "Workshop on Science Data Management at the National Institute of Polar Research"

Mitsuo Fukuchi[1], Lee Belbin[2], David Watts[2] and Toru Hirawake[1]

国立極地研究所のサイエンス・データマネージメントに関するワークショップ報告

福地光男[1]・Lee Belbin[2]・David Watts[2]・平譯　享[1]

(Received October 21, 2004; Accepted January 20, 2005)

要旨:　2004年2月25–27日に国立極地研究所にて,「国立極地研究所のサイエンス・データマネージメントに関するワークショップ」が開催された. 南極研究科学委員会と南極観測実施責任者評議会により設立された「合同南極データマネージメント委員会」において主導的な活動をしてきたオーストラリア南極局のデータマネージャー等2名を招待し, 南極局におけるおよそ過去10年におけるデータマネージメントの発展をレビューし, 国立極地研究所におけるデータマネージメントの今後の方向性について討議した.

Abstract: The Workshop on Science Data Management was held at the National Institute of Polar Research (NIPR) from February 25–27, 2004. The Manager and Senior Applications developer from the Australian Antarctic Data Centre (AADC) were invited to distil the development and operation of the AADC in the context of Antarctic science data management and the Joint Committee on Antarctic Data Management (JCADM). The current data management situation and future requirements at NIPR were identified.

1. Introduction

Dr. Mitsuo Fukuchi (Director, Center for Antarctic Environment Monitoring, National Institute of Polar Research) invited Lee Belbin, the Manager of the Australian Antarctic Data Centre and David Watts (Senior Application Developer) to Tokyo in February 2004 to discuss options for science data management at the National Institute for Polar Research. Lee established the Australian Antarctic Data Centre (AADC) in 1995 and had chaired the first three years of the SCAR/COMNAP Joint Committee on Antarctic Data Management (JCADM: http://www.jcadm.scar.org). The AADC now has ten staff providing a broad range of services to all areas of the Australian Antarctic Program. The program and the participants at the workshop are listed in Tables 1 and 2, respectively.

[1] 国立極地研究所. National Institute of Polar Research, Research Organization of Information Systems, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515.

[2] Australian Antarctic Division, Channel Highway, Kingston, Tasmania 7050, Australia.

Table 1. Program of "Workshop on Science Data Management at the National Institute of Polar Research (NIPR)".

Program of Workshop on Data Management

25–26 February 2004 at Lecture Room, NIPR

Day 1: February 25 (Wed.) Introduction day

10:30–10:45 Introduction by M. Fukuchi (Director of Center for Antarctic Environment Monitoring)

Structure and function of NIPR, which was established in 1973, from a viewpoint of data management

10:45–12:00 Development of Data Management at Australian Antarctic Division by L. Belbin

12:00–13:30 Lunch break

13:30–15:30 Role of research center at NIPR

Information Science Center founded in 1990 by N. Sato (Director of Information Science Center)

Arctic Environment Research Center in 1990 by Y. Fujii (Director of Arctic Environment Research Center)

Center for Antarctic Environment Monitoring in 1995 by M. Fukuchi (Director of Center for Antarctic Environment Monitoring)

Antarctic Meteorite Research Center in 1998 by K. Shiraishi (Director of Antarctic Meteorite Research Center)

15:30–16:00 Work plan arrangement for Day 2

Day 2: February 26 (Thurs.) Technical day

10:30–12:00 Technical introduction by D. Watts (AAD)

12:00–13:30 Lunch break

13:30–15:30 Technical discussion on data management

Day 3: Wrap up for future development

10:30–12:00 Discussion on suggestion and comment by L. Belbin

Table 2. Participant list of "Workshop on Science Data Management at the National Institute of Polar Research" held in February 25–27, 2004.

Name of participants	Affiliation
Lee Belbin	Australian Antarctic Division
David Watts	Australian Antarctic Division
Mitsuo Fukuchi	National Institute of Polar Research
Takashi Yamanouchi	National Institute of Polar Research
Kazuo Shibuya	National Institute of Polar Research
Makoto Taguchi	National Institute of Polar Research
Toru Hirawake	National Institute of Polar Research
Yoshiyuki Fujii	National Institute of Polar Research
Hiroshi Kanda	National Institute of Polar Research
Natsuo Sato	National Institute of Polar Research
Kazuyuki Shiraishi	National Institute of Polar Research
Masaki Kanao	National Institute of Polar Research

2. Scope of SCAR/COMNAP JCADM

The JCADM group was established by SCAR and COMNAP to address Antarctic science data management issues. The committee's first priority was to encourage and assist Antarctic Treaty nations involved in Antarctic research to establish their National Antarctic Data Centers (NADCs). Each NADC would be established in a form that would best fit the nature of the nation's Antarctic science activity. For example, Australia decided to combine science data management, Antarctic mapping and state of the Antarctic environment reporting into their data center functions. Other NADCs focused on data management of the nation's science priority disciplines. Some NADCs limited their activity only to the creation of metadata.

JCADM encouraged science administrators to attend at least one JCADM meeting before appointing someone to manage or run their NADC. This strategy enabled nations to better select the type of person needed to lead in the management of their science data requirements. For the first five years, JCADM encouraged emerging NADCs to focus on metadata. Metadata is a standardized description of data. Metadata extends an index like a library catalogue to include parameters that would aid discovery and use of the data. Metadata parameters include author, title, location and time of the data collection, data format, data usage constraints and keywords. A metadata catalogue such as the Antarctic Master Directory (http://gcmd.gsfc.nasa.gov/Data/portals/amd/) can be searched by free text or metadata parameters.

The emphasis on metadata would in JCADM's opinion, ensure that new data would be catalogued to international standards in an international directory and offer greatest value to Antarctic science in addressing Article III.1.c. of the Treaty which states that "scientific information should be fully and freely exchanged." Valuable scientific data would be preserved and accessible for cooperation and collaboration into the future.

3. The Objectives of workshop at NIPR

Lee Belbin and David Watts were invited to a workshop on data management February 25–27, 2004 at NIPR to share their experiences from the Australian Antarctic Data Centre (AADC: http://www.aad.gov.au/default.asp?casid=3786). There were four components to the workshop—

1) An extended presentation by Lee Belbin on the 'AADC Success Story' to NIPR Science Program Leaders and other interested scientists (see Fig. 1),
2) Presentations by NIPR science program leaders on the nature of their research,
3) A presentation by David Watts on core technical infrastructure issues in science data management (Appendix 1),
4) A discussion by Lee Belbin combining responses to key questions about data management posed by Mitsuo Fukuchi and Lee Belbin's observations on the NIPR data management position (Fig. 2).

Figure 1 was generated By Lee Belbin as a response to Mitsuo Fukuchi's question "Why is the AADC so successful?" while Fig. 2 answered six questions submitted by Mitsuo that are fundamental to the establishment of an effective data management strategy for NIPR. Figures 1 and 2 are complementary. Some overlap in the information in these Figs. 1 and 2 provides additional emphasis on the most significant issues that need to be considered by NIPR in establishing a data management strategy. For example, in Fig. 1, Lee Belbin's acknowledgement of the importance of the environment in which the AADC was established aligns with answers to question 1 in Fig. 2 "How was the AADC established and developed?"

4. Establishment and development of Australian Antarctic Data Centre

The AADC was established by Australia to preserve Australia's Antarctic science data, to address Article III.1.c. of the Antarctic Treaty and to fulfill Australian government goals on spatial data infrastructure. The AADC was established with two staff members, a manager (Lee Belbin) and a Mapping Officer (Henk Brolsma). Over the past nine years, the AADC has employed up to 15 staff members and currently has nine 'ongoing' positions and one contract position. This growth would not have occurred unless the AADC was seen to be providing a cost-effective service to the Australian Antarctic Program.

The AADC has been strategic in building an innovative infrastructure and a range of effective applications. For example, a Web-based research proposal system was written to capture scientific research project information at the time of submission by principal investigators. Using this strategy, metadata could be automatically generated from proposal content without the need for scientists to re-enter basic information. Lee was also responsible for ATCM XII Resolution 4 (1988: http://www.jcadm.scar.org/TreatyDocs/ATCM_98_resolution.htm) promoting NADC establishment and metadata priorities. This resolution prompted Australia to develop an Antarctic data management policy (http://www.aad.gov.au/default.asp?casid=3959) that sets the foundation for science data management for Australia's Antarctic Science Program. This policy was endorsed by Australia's peak Antarctic science committee, the Antarctic Science Advisory Committee. The data management policy stipulates that all projects must submit data to the AADC within two years of data collection. These data publicly available online and are linked to the online metadata system (http://www.aad.gov.au/default.asp?casid=3802).

The AADC receives data, checks the consistency and quality of the metadata, and makes the data freely available online. The centre also provides a wide range of value-added services. Where feasible, datasets are combined into Web-accessible databases (http://www.aad.gov.au/default.asp?casid=3803). Such databases simplify searching and subsetting of data by the science community. The AADC maintains over 30 such databases covering publications, biodiversity, meteorology, oceanography, events and maps among others.

The AADC also manages Australia's Antarctic Mapping Program, provides advice on data management, GIS, mapping (including global positioning systems), data analysis and drives Australia's Antarctic state of the environment reporting system (SIMR: http://www.aad.gov.au/default.asp?casid=3808). During interactions with scientists, the centre also gathers information to create a series of public educational pages on the Web (http://www.aad.gov.au/default.asp?casid=3249).

Figure 1 provides an outline of what we believe are the significant factors contributing to the success of the AADC. The diagram is an example of what is termed a 'Mind Map' (Buzan, 1993); a tree structure that displays relationships between any set of objects. The map in Fig. 1 was prepared prior to the workshop and refined during the workshop to ensure that key issues raised by NIPR program leaders were addressed. The 'map' attempts to structure success into a series of headings such as the three phases of development of the centre (establishment, current and future prospects). Lower order connections on the diagram provide the answers to questions. For example, support from senior management, demonstrated leadership and management skills, and the right staff were identified as important factors leading to the acceptance of the AADC as a vital component of an effective Antarctic research program.

Figure 2 uses the same structure to provide answers to specific questions on science data management posed before the workshop. For example, question 1 asks "How was the AADC established and developed?" This map was developed during the workshop as an understanding of the data management situation at NIPR emerged. Important components of this map included an obligation to maintain an effective repository of very expensive data, cost-benefits to research by reducing duplication and simplifying access to data and accountability.

5. Key data management issues identified through the workshop

The recognition that while the priority for NIPR is research, output and outcomes may be enhanced through developing a data management infrastructure. Such an infrastructure would enable NIPR to adapt to a changing political environment, would assist program leaders in managing research projects and assist scientists in locating and re-using valuable Antarctic data.

At NIPR, scientists are currently responsible for their own data management. Data management is mainly associated with desktop applications such as Excel and specialized analytical applications such as statistical packages. Generalized data repositories are rare, and when they do exist, are limited to a few desktop computers, rather than being widely available through the Web. At NIPR, the management of, and access to data is dependent on a few key people. The natural outcome of such a strategy is inevitable loss of valuable data, reduction in research time, and lack of a comprehensive and systematic knowledge of science outputs and outcomes.

The most important factors leading to the success of the AADC were Web-accessible metadata, and the project and publications databases. These three applications provided the 'backbone' of science data management within the Australian Antarctic Program. Linkages between these databases enabled information to be tracked from project initiation to project completion. Management and reporting on these databases requires only a Web browser, and follows the basic principle "*store once—use many times for many different applications*". A demonstration Web site http://aadc-maps.aad.gov.au/aadc/nipr/ has been developed by the AADC for the NIPR. This Web site includes test databases on science projects, publications and metadata. Maintenance and documentation is also included on this site.

Data management infrastructure includes policy and procedures, not just hardware and software. Without an effective Antarctic science data management policy, the AADC would not have been successful. This policy states that scientists have two years exclusive use of their data, unless a good case can be made to the Chief Scientist to extend this period. After two years, data must be documented with quality metadata and given to the AADC. The AADC checks data and metadata consistency and places the data online for public access. Value is added to the data by ensuring that all variables are recorded in standard units (for example, all temperature data is in degrees centigrade) and described by a central data dictionary. This strategy greatly simplifies the creation of composite databases when data reaches a critical mass. This strategy has also enabled effective data mining of the repository to occur; new relationship to be detected, errors identified and research targeted.

Reference

Buzan, T. (1993): The Mind Map Book. London, BBC Books, 320 p.

Appendix 1. Technical issue for data management.

Infrastructure

From what we understand about the NIPR environment, we would recommend the use of either Sun Solaris on Sun hardware or Linux on an Intel or equivalent hardware platform. This combination is highly stable and secure.

Application Environment

Most web site problems can be attributed to an overly complex application environment. We believe there are two choices for application hosting and development, J2EE server or Microsoft.Net. These are both mature technologies with most open-source projects based on the J2EE platform. The Data Centre has selected the Java-based application ColdFusion to develop its web-enabled databases. Being based on Java, it can run on any platform unlike any.Net product which is restricted to Windows OS.

The Data Centre uses ColdFusion because it is a fast, complete web scripting solution. It has a simple learning curve and hides underlying complexity such as database access. It can access multiple databases (SQL, Access, Oracle etc) providing a seamless experience for the end user. Data can be migrated from database to database without a user being aware.

Site Philosophy and Design

It is important for the end-user to experience a site with a common look and feel. Currently, the NIPR web site is a mixture of navigation and terminology making discovery of resources difficult. COMNAP provides an example of a simple and consistent site. Once the site is established the developer can manage content without spending unnecessary time on 'cosmetics'. The AADC uses a combination of simple search mechanisms for rapid responses and more complex search mechanisms for advanced users. Most of the AADC databases have a common database interface therefore users experience a consistent 'look and feel'.

Database Design

The AADC has developed a database of all known AAD databases. A public list can be seen at http://aadc-maps.aad.gov.au. Where possible, all database design parameters are stored in this database and the web pages are automatically derived from this content. To change web pages requires only minimal database editing.

Controlled parameters and keywords are fundamental requirements for efficient database design. These features enable the use of ontologies and cross-linking of data from various databases. An AADC example is the Antarctic artifacts database, containing 300 keywords with complete descriptions (see http://aadc-maps.aad.gov.au/aadc/artefacts/).

The AADC uses three types of data cross-links—

1. Explicit links—*e.g.* Link Map reference 12345 to Taxa reference 78. The reverse lookup from taxa to map is automatically shown. For example http://aadc-maps.aad.gov.au/aadc/gaz/display_name.cfm?gaz_id=50039
2. Link via web page parameters with position extents and/or date ranges. For example, a user has found a map and wishes to list any species observed within the map bounds.
3. Text matches across multiple databases. For example, search for "Mawson".

Fundamental Databases

Projects

The Science Project is the fundamental research unit. The AAD project database is complex but a simpler version could be established based on—

- Project Number or Code (*e.g.* CAEM-1)
- Project Title
- Investigator—include contact details
- Objectives/Aims
- Where is the work to be done - on a seasonal basis *e.g.* 2003/04 Greenland
- Which program area (*e.g.* CAEM)
- Project status per season—*e.g.* New, Approved, Withdrawn, Rejected

The project details can be used to create a preliminary metadata record that can be updated after field work and analysis. The metadata record can be globally discovered and provide links to the project.

Publications

Scientific productivity has been determined largely by scientific publications, but we would advocate value to scientific datasets and metadata. The AAD maintains a publications database that is also cross-linked to the projects database. A simple report for a project can list all its outcomes; publications, metadata and data.

Metadata

The Global Change Master Directory of NASA (http://gcmd.nasa.gov/) provides a metadata hosting service for the NADCs. The AADC hosts its own metadata database but we would recommend that NIPR uses the GCMD to host their metadata but to ensure that the links from projects to metadata are maintained via URL's. Storing metadata records at NIPR would substantially increase the complexity of the site and could perhaps be considered in future developments.

A preliminary demonstration of a project-publications-metadata application has been written in ColdFusion to demonstrate some of the proposed functionality (see http://aadc-maps.aad.gov.au/aadc/nipr/).

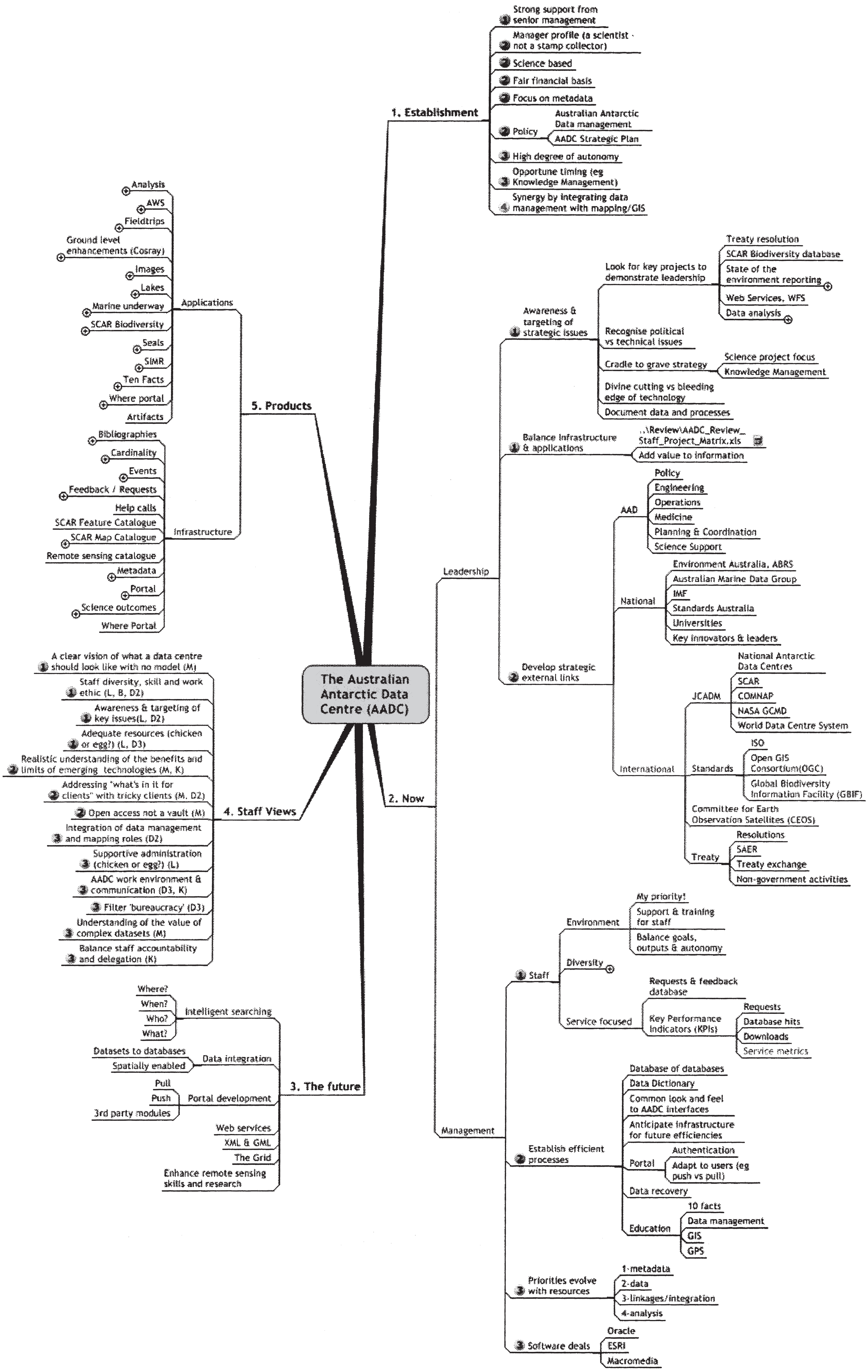
The Australian Antarctic Data Centre (AADC)
1. Establishment
Strong support from senior management
Manager profile (a scientist - not a stamp collector)
Science based
Fair financial basis
Focus on metadata
Policy
Australian Antarctic Data management
AADC Strategic Plan
High degree of autonomy
Opportune timing (eg Knowledge Management)
Synergy by integrating data management with mapping/GIS
2. Now
Leadership
Awareness & targeting of strategic issues
Look for key projects to demonstrate leadership
Treaty resolution
SCAR Biodiversity database
State of the environment reporting
Web Services, WFS
Data analysis
Recognise political vs technical issues
Cradle to grave strategy
Science project focus
Knowledge Management
Divine cutting vs bleeding edge of technology
Document data and processes
Balance infrastructure & applications
..\Review\AADC_Review_Staff_Project_Matrix.xls
Add value to information
Develop strategic external links
AAD
Policy
Engineering
Operations
Medicine
Planning & Coordination
Science Support
National
Environment Australia, ABRS
Australian Marine Data Group
IMF
Standards Australia
Universities
Key innovators & leaders
International
JCADM
National Antarctic Data Centres
SCAR
COMNAP
NASA GCMD
World Data Centre System
Standards
ISO
Open GIS Consortium(OGC)
Global Biodiversity Information Facility (GBIF)
Committee for Earth Observation Satellites (CEOS)
Treaty
Resolutions
SAER
Treaty exchange
Non-government activities
Management
Staff
Environment
My priority!
Support & training for staff
Balance goals, outputs & autonomy
Diversity
Service focused
Requests & feedback database
Key Performance Indicators (KPIs)
Requests
Database hits
Downloads
Service metrics
Establish efficient processes
Database of databases
Data Dictionary
Common look and feel to AADC interfaces
Anticipate infrastructure for future efficiencies
Portal
Authentication
Adapt to users (eg push vs pull)
Data recovery
Education
10 facts
Data management
GIS
GPS
Priorities evolve with resources
1-metadata
2-data
3-linkages/integration
4-analysis
Software deals
Oracle
ESRI
Macromedia
3. The future
Intelligent searching
Where?
When?
Who?
What?
Data integration
Datasets to databases
Spatially enabled
Portal development
Pull
Push
3rd party modules
Web services
XML & GML
The Grid
Enhance remote sensing skills and research
4. Staff Views
A clear vision of what a data centre should look like with no model (M)
Staff diversity, skill and work ethic (L, B, D2)
Awareness & targeting of key issues(L, D2)
Adequate resources (chicken or egg?) (L, D3)
Realistic understanding of the benefits and limits of emerging technologies (M, K)
Addressing "what's in it for clients" with tricky clients (M, D2)
Open access not a vault (M)
Integration of data management and mapping roles (D2)
Supportive administration (chicken or egg?) (L)
AADC work environment & communication (D3, K)
Filter 'bureaucracy' (D3)
Understanding of the value of complex datasets (M)
Balance staff accountability and delegation (K)
5. Products
Applications
Analysis
AWS
Fieldtrips
Ground level enhancements (Cosray)
Images
Lakes
Marine underway
SCAR Biodiversity
Seals
SIMR
Ten Facts
Where portal
Artifacts
Infrastructure
Bibliographies
Cardinality
Events
Feedback / Requests
Help calls
SCAR Feature Catalogue
SCAR Map Catalogue
Remote sensing catalogue
Metadata
Portal
Science outcomes
Where Portal

Fig. 1. *This figure attempts to answer the question as to why the Australian Antarctic Data Centre (AADC) has been successful. A significance level (where 1 is the highest and 4 is the lowest) has been assigned to the main establishment, current situation and the staff view issues. For example, the most important factor in the establishment of the AADC was that there was strong support from senior Antarctic Division management for the establishment of the AAD. Of only slightly less significance (priority 2) were the profile of the AADC Manager, that the AADC was based in the Science Program, that there was a fair allocation of funds, an emphasis on metadata and basic policy.*

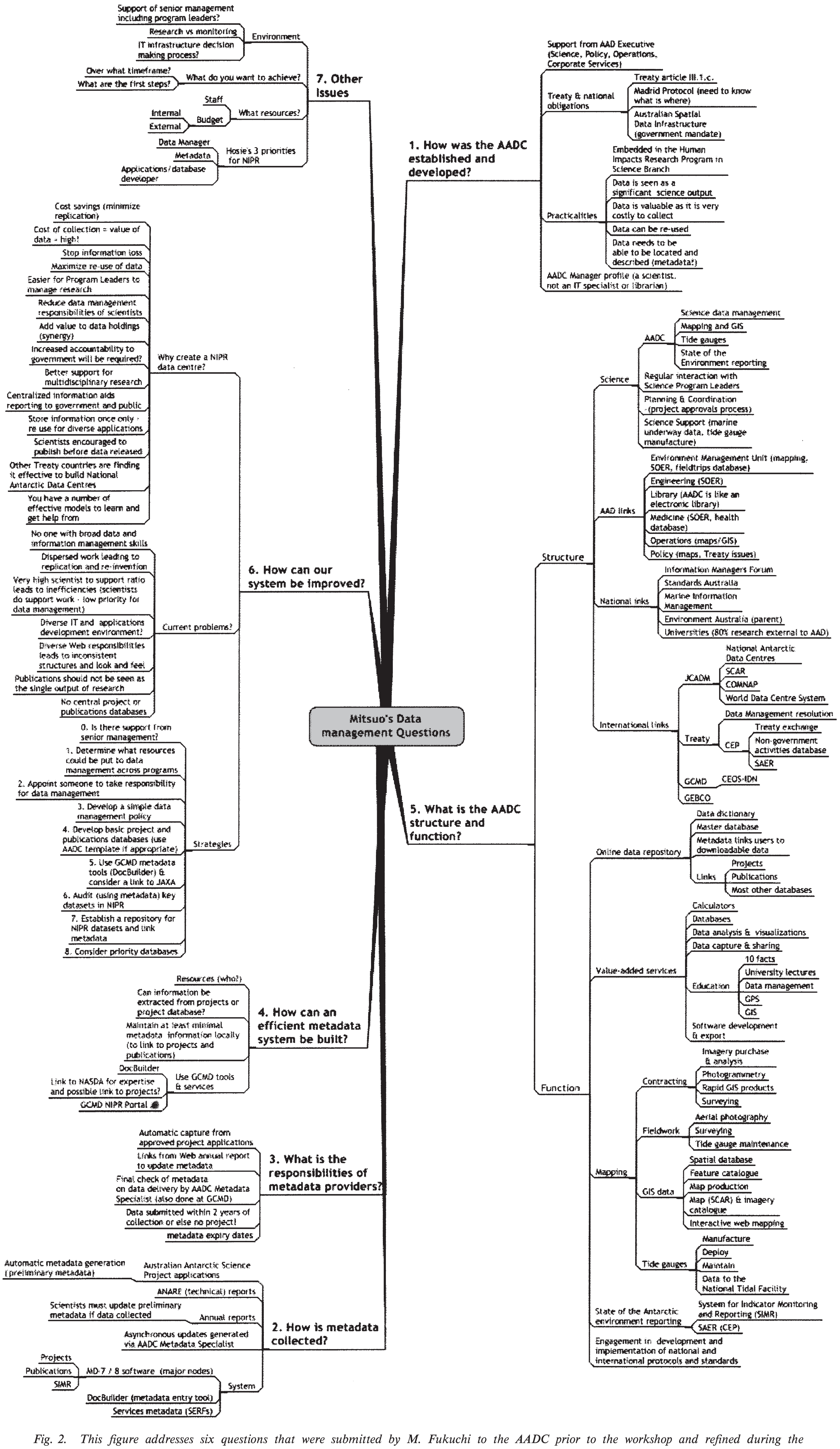
Mitsuo's Data management Questions
1. How was the AADC established and developed?
Support from AAD Executive (Science, Policy, Operations, Corporate Services)
Treaty & national obligations
Treaty article III.1.c.
Madrid Protocol (need to know what is where)
Australian Spatial Data Infrastructure (government mandate)
Embedded in the Human Impacts Research Program in Science Branch
Practicalities
Data is seen as a significant science output
Data is valuable as it is very costly to collect
Data can be re-used
Data needs to be able to be located and described (metadata!)
AADC Manager profile (a scientist, not an IT specialist or librarian)
5. What is the AADC structure and function?
Structure
Science
AADC
Science data management
Mapping and GIS
Tide gauges
State of the Environment reporting
Regular interaction with Science Program Leaders
Planning & Coordination (project approvals process)
Science Support (marine underway data, tide gauge manufacture)
AAD links
Environment Management Unit (mapping, SOER, fieldtrips database)
Engineering (SOER)
Library (AADC is like an electronic library)
Medicine (SOER, health database)
Operations (maps/GIS)
Policy (maps, Treaty issues)
National inks
Information Managers Forum
Standards Australia
Marine Information Management
Environment Australia (parent)
Universities (80% research external to AAD)
International links
JCADM
National Antarctic Data Centres
SCAR
COMNAP
World Data Centre System
Treaty
Data Management resolution
CEP
Treaty exchange
Non-government activities database
SAER
GCMD
CEOS-IDN
GEBCO
Function
Online data repository
Data dictionary
Master database
Metadata links users to downloadable data
Links
Projects
Publications
Most other databases
Value-added services
Calculators
Databases
Data analysis & visualizations
Data capture & sharing
Education
10 facts
University lectures
Data management
GPS
GIS
Software development & export
Mapping
Contracting
Imagery purchase & analysis
Photogrammetry
Rapid GIS products
Surveying
Fieldwork
Aerial photography
Surveying
Tide gauge maintenance
GIS data
Spatial database
Feature catalogue
Map production
Map (SCAR) & imagery catalogue
Interactive web mapping
Tide gauges
Manufacture
Deploy
Maintain
Data to the National Tidal Facility
State of the Antarctic environment reporting
System for Indicator Monitoring and Reporting (SIMR)
SAER (CEP)
Engagement in development and implementation of national and international protocols and standards
7. Other Issues
Environment
Support of senior management including program leaders?
Research vs monitoring
IT infrastructure decision making process?
What do you want to achieve?
Over what timeframe?
What are the first steps?
What resources?
Staff
Budget
Internal
External
Hosie's 3 priorities for NIPR
Data Manager
Metadata
Applications/database developer
6. How can our system be improved?
Why create a NIPR data centre?
Cost savings (minimize replication)
Cost of collection = value of data - high!
Stop information loss
Maximize re-use of data
Easier for Program Leaders to manage research
Reduce data management responsibilities of scientists
Add value to data holdings (synergy)
Increased accountability to government will be required?
Better support for multidisciplinary research
Centralized information aids reporting to government and public
Store information once only - re use for diverse applications
Scientists encouraged to publish before data released
Other Treaty countries are finding it effective to build National Antarctic Data Centres
You have a number of effective models to learn and get help from
Current problems?
No one with broad data and information management skills
Dispersed work leading to replication and re-invention
Very high scientist to support ratio leads to inefficiencies (scientists do support work - low priority for data management)
Diverse IT and applications development environment?
Diverse Web responsibilities leads to inconsistent structures and look and feel
Publications should not be seen as the single output of research
No central project or publications databases
Strategies
0. Is there support from senior management?
1. Determine what resources could be put to data management across programs
2. Appoint someone to take responsibility for data management
3. Develop a simple data management policy
4. Develop basic project and publications databases (use AADC template if appropriate)
5. Use GCMD metadata tools (DocBuilder) & consider a link to JAXA
6. Audit (using metadata) key datasets in NIPR
7. Establish a repository for NIPR datasets and link metadata
8. Consider priority databases
4. How can an efficient metadata system be built?
Resources (who?)
Can information be extracted from projects or project database?
Maintain at least minimal metadata information locally (to link to projects and publications)
Use GCMD tools & services
DocBuilder
Link to NASDA for expertise and possible link to projects?
GCMD NIPR Portal
3. What is the responsibilities of metadata providers?
Automatic capture from approved project applications
Links from Web annual report to update metadata
Final check of metadata on data delivery by AADC Metadata Specialist (also done at GCMD)
Data submitted within 2 years of collection or else no project!
metadata expiry dates
2. How is metadata collected?
Australian Antarctic Science Project applications
Automatic metadata generation (preliminary metadata)
ANARE (technical) reports
Annual reports
Scientists must update preliminary metadata if data collected
Asynchronous updates generated via AADC Metadata Specialist
System
MD 7 / 8 software (major nodes)
Projects
Publications
SIMR
DocBuilder (metadata entry tool)
Services metadata (SERFs)

Fig. 2. *This figure addresses six questions that were submitted by M. Fukuchi to the AADC prior to the workshop and refined during the workshop. Four other factors are also included. The issues identified in this figure and Fig. 1 provide a strategy for the establishment of an effective data management function at NIPR.*