



# BCIS

---

**BIODIVERSITY CONSERVATION  
INFORMATION SYSTEM**

**Framework for Information Sharing  
Series Editor John R. Busby**

**VOLUME 1:  
PRINCIPLES**

---

©2000. Biodiversity Conservation Information System. Reproduction of this publication for educational or other non-commercial purposes is authorised without prior permission from the copyright holders, provided the source is acknowledged. Reproduction for resale or other commercial purpose is prohibited without the prior written permission of the copyright holders.

The views expressed in this handbook do not necessarily reflect those of individual BCIS Members or Partners. The designations of geographical entities in this report and the presentation of the material do not imply the expression of any opinion whatsoever on the part of BCIS or its Members, or other participating organisations concerning the legal status of any country, territory, or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Citation: Biodiversity Conservation Information System. 2000. *Framework for Information Sharing: Principles* . Busby, J.R. (Series Editor).

Available from Program Manager, BCIS [contact details on <http://www.biodiversity.org>]

*Every reasonable care has been taken to check that all web addresses are correct at the time of writing (December 1999). Regrettably, sites and pages change over time so some addresses may no longer be correct.*

# VOL. 1: PRINCIPLES ♦ CONTENTS

---

<b>Acknowledgements</b> .....	5
<b>BCIS Framework for Information Sharing</b> .....	6
<b>Manager's Guide</b> .....	8
Context .....	8
Actions .....	9
<b>1. The Need for Information Infrastructure</b> .....	11
1.1 Introduction .....	11
1.2 The BCIS Information Management Framework .....	11
1.3 Data and Information—The Need for a Policy .....	12
1.4 Data Management in Context .....	13
1.5 The Significance of a Data Policy .....	14
1.6 National and International Implications and Related Activities .....	15
1.7 Commercial Potential .....	16
1.8 The Data—Information Relationship .....	16
<b>2. General Principles</b> .....	17
2.1 Introduction .....	17
2.2 Goal .....	17
2.3 Strategies .....	18
2.4 The Data Management Process .....	19
<b>3. Proposed Structure</b> .....	20
3.1 Introduction .....	20
3.2 Co-ordination .....	20
3.2.1 Introduction .....	20

3.2.2	Data Management Group	20
3.2.2.1	Background	20
3.2.2.2	Proposed Structure	21
3.3	Data Management Infrastructure	21
3.3.1	Introduction	21
3.3.2	Data Collectors	22
3.3.3	Data Owners	22
3.3.4	Data Custodians	23
3.3.4.1	Introduction	23
3.3.4.2	Data Custodian Responsibilities and Obligations	24
3.3.5	Information Agencies	24
3.3.6	Physical Infrastructure	25
3.3.6.1	Electronic Networks	25
3.3.6.2	Technology	25
3.4	Core Datasets	26
3.5	Analysis and Dissemination Tools	27
3.6	Standards, Guidelines and Procedures	27
3.6.1	Introduction	27
3.6.2	Quality Assurance	28
3.6.3	Standards	28
3.6.4	Metadata	28
3.6.5	Data Exchange	28
<b>4.</b>	<b>Implementation Challenges</b>	<b>30</b>
4.1	Freeing up the Flow of Data	30
4.2	Resources for Data Management	31
4.3	Resolving Ownership and Custodianship	31
	<b>Further Reading</b>	<b>32</b>

## ACKNOWLEDGEMENTS

---

These handbooks have been loosely modelled on the *WCMC Handbooks on Biodiversity Information Management*, edited by Jake Reynolds of the World Conservation Monitoring Centre and published by the Commonwealth Secretariat, London. These, in turn, benefited from experiences gained through the Biodiversity Data Management (BDM) Project, administered by the United Nations Environment Programme (UNEP) and funded by the Global Environment Facility (GEF), and related initiatives supported through the European Union (EU) and European Environment Agency (EEA).

Fundamental to the development of this series have been the contributions from numerous colleagues working in the field of biodiversity information management. Among these, particular mention goes to Dr Jake Reynolds and other staff of WCMC; Professor Ian Crain and Gwynneth Martin of the Orbis Institute, Ottawa; Kevin Grose, former Head of the Information Management Group at IUCN–The World Conservation Union; and staff of the Environmental Resources Information Network (ERIN) Unit, Environment Australia, and of the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia.

This Handbook has been adapted from *National Marine Data Policy: Discussion Paper*, by B. Searle, Head, Australian Oceanographic Data Centre, June 1997.

---

*The BCIS consortium Members extend their thanks and gratitude to the Norwegian Agency for Development Cooperation (NORAD) for providing financial support to develop the BCIS concept and these handbooks.*

# **BCIS FRAMEWORK FOR INFORMATION SHARING**

---

The purpose of the *BCIS Framework for Information Sharing* is to support BCIS Members and others making decisions on the conservation and sustainable use of living resources. The handbooks form part of a comprehensive set of supporting materials designed to build information-management capacity and improve decision-making.

The intended audience includes senior managers in Member organisations, their equivalents in other organisations, information and environmental-science professionals, and others who have a stake in the use or management of living resources. Although written to address the specific need for improved management of biodiversity-related information within the BCIS network, the underlying principles apply to environmental information networks in general, and to decision-making at all levels. The issues and concepts presented may also be applied in sectors other than biodiversity conservation: forestry, agriculture, wildlife management and beyond.

The handbooks deal with a range of issues and processes relevant to the use of information in decision-making, including the strengthening of organisations and organisational linkages, data custodianship and management, metadata and the development of infrastructure to support data and information exchange. Experience suggests that some of the greatest challenges in information management today are concerned with organisational issues, rather than technical or scientific concerns. Consequently, topics are addressed at management and organisational levels, rather than from a technical or scientific standpoint. Nevertheless, in adopting this framework approach, BCIS has tried to adhere to recognised conventions and formalisms used in information management.

Overall, the handbook series comprises:

## Executive Overview

- **Volume 1: Principles**
- Volume 2: Procedures Manual
- Volume 3: Custodianship
- Volume 4: Data Access
- Volume 5: Metadata
- Volume 6: Standards & Quality Assurance
- Volume 7: Core Datasets
- Volume 8: Tools & Technologies

Collectively, the handbook series promotes a shift from tactically based information systems, aimed at supporting individual projects, to strategic systems that promote the development of information infrastructure through the building of capacity within BCIS and other networks. This approach not only encourages data to be managed more effectively within organisations, but also encourages data to be shared amongst organisations for the development of the integrated products and services needed to address complex and far-reaching environmental issues.

The handbook series can be used in a number of ways. Individual handbooks can be used to guide managers and professional staff on specific aspects of information management or they can be used collectively as a reference source for strategic planning and project development.

This Handbook makes the case for investment in strategic data and information infrastructure as a fundamental corporate resource. It highlights the value of mobilising scientific information, within collaborative interdisciplinary networks, for policy-development and decision-support. It also introduces the need for a data policy to guide information infrastructure development.

# MANAGER'S GUIDE

---

## Context

Strategically planned, professionally organised information infrastructure is vital for effectively fulfilling BCIS objectives. Individual Members are custodians of vast amounts of information, which is collected, managed, processed and disseminated at considerable cost. Data and information management and use is a major responsibility of each agency, and not simply a by-product of an activity or service. Information is a strategic resource of exceptional value, not only to the individual Member, but also to the other Members and other sectors. While this information is valuable, many potential users within and outside BCIS either do not know that it exists or do not know how to use it. Improved visibility and accessibility of the information will promote opportunities for development of value-added services, which will more than justify the investments made.

Policies and procedures are needed to guide the transition from tactical, project-based data collection and management to strategic information infrastructure that will inform decision-making on a wide range of issues and at a wide range of scales. Currently data are, with some exceptions, inadequate, inconsistent and inaccessible. Putting data to use outside the immediate discipline of origin can be very difficult. Comparatively few data are available within timeframes required to support operational needs, particularly emergency response. Present users of data are faced with a confusing array of datasets and data formats and the lack of applied standards results in high data integration and interpretation costs. Overlapping and confused data custodianship and corporate mandates result in further costs from duplication of effort, leading to poor decisions and lost opportunities. Wider understanding of key data management principles and the implementation of a data policy is the answer.

BCIS data-management principles provide the framework within which the activities of data collectors, data managers and information providers can be integrated. They scope a distributed infrastructure and sets of standards, guidelines and procedures to improve the effectiveness of data management and the creation and dissemination of priority information in support of conservation, sustainable development and related issues.

The potential value of data to each agency is very substantial. However, in general, managers give a low priority to its collection and management. Currently, integrated data and information systems are so underdeveloped that the available information cannot be readily applied to solve environmental questions at any scale, from local to global.

Data management will only be given the necessary priority if management sees tangible direct benefits outweighing the costs. These benefits of investing in better-managed data include enhanced speed and cost-effectiveness in:

- ingesting standardised data from external sources;
- integrating external with internal data;
- generating more comprehensive, and thus more attractive, useful and valuable information products; and
- mobilising data for others, under collaborative arrangements or for some public-good purpose.

### **Actions**

Managers should focus on the need to achieve the data management goal of: “An integrated and technologically advanced distributed data management, data and information processing and dissemination capability to support data and information needs of decision-makers in developing countries and implementing environmentally sustainable development policies at all levels.”

The goal will be achieved through the formalisation of infrastructure, creation of guidelines, standards and procedures to support data management and processing and the generation and distribution of data and information products. This process will be implemented by a well-supported, science-based and technically competent distributed group of data custodians and related agencies.

Managers should facilitate the development and implementation of a data policy, which could encompass the following strategies to achieve the goal:

- create an easily accessible, distributed data network to manage and disseminate data collected as part of project and other activities in support of corporate objectives;
- develop core data sets as baseline products;
- provide ‘best practice’ quality-assurance mechanisms and procedures to create validated, well-documented data sets to meet priority information needs;
- archive, where appropriate, all data collected, to ensure availability of the data for multiple use and to safeguard it for future generations;

- improve the cost-effectiveness of agency policy and program development through the coordination of data and information activities;
- provide timely and operationally significant data and information products to support a range of activities, with an emphasis on supporting all sectors of society in developing environmentally sustainable development and economic processes;
- make data available to the widest range of users, at the highest quality and in as timely fashion as practicable; and
- provide information services and products in the public interest.

A data policy covers the following main elements to underpin the above strategies:

- a coordination mechanism;
- the physical infrastructure of agencies and partner organisations that make up a network;
- core datasets;
- analysis and dissemination tools; and
- the standards, guidelines and procedures by which the relevant agencies operate.

All this will require the adoption and adaptation of best practice in professional data and information management for data collectors, data owners, data custodians and information-generating agencies. This Handbook specifies how agencies should implement the above strategies and thus achieve their goal.

# 1. THE NEED FOR INFORMATION INFRASTRUCTURE

---

## 1.1 Introduction

Governments, agencies and individuals are increasingly acknowledging the significance of data and information in:

- supporting decision-making processes at all levels;
- assisting the development of environmental policies;
- improving the effectiveness of management programmes; and
- generally expediting progress towards environmental sustainability.

The decision of 12 international non-governmental organisations and networks to work more closely together within the framework of the Biodiversity Conservation Information System (BCIS) signals the increasing emphasis being placed on information as a strategic resource. Relevant data and information should be marshalled and made readily available to other agencies, governments, the corporate sector and civil society to support and better-inform decisions. Increasingly, environmental information is viewed as a crucial component within the principles of conservation and sustainable development.

One example of the strategic use of such information is the Pacific Forest Plan for the forests of the Pacific Northwest of the United States. This Plan focused on “providing [for] biological diversity and sustainable economic activity across the entire landscape. The final landscape plan, mapped out by teams of scientists from myriad sources, and constructed bit-by-bit using modern geographic information technology, has no precedent on this scale. It marks the beginning of a new, applied science of landscape conservation, using the tools of many sciences to find an acceptable balance between human communities and nature.” (Babbitt, 1995).

## 1.2 The BCIS Information Management Framework

The BCIS Steering Committee, comprising managerial and professional staff representing the members, has clearly identified the significance of information in effectively fulfilling BCIS objectives. It recognises that the Members are custodians of vast amounts of information, which is collected, managed, processed and disseminated at considerable cost. Information is a major responsibility of each agency, and not simply a by-product of an activity or service. This information is a strategic resource of exceptional value, not only to the individual Member, but

also to the other Members, the non-governmental, government and corporate sectors, and civil society. Unfortunately, while this information is valuable, many potential users within and outside BCIS either do not know that it exists or do not know how to use it. Improved visibility and accessibility of the information stored within agencies (other than material protected by privacy or confidentiality provisions) is beneficial in itself. But it will also promote increased opportunities for partnerships and for development of value-added services, which will leverage the investment made in collecting the information in the first place.

### **1.3 Data and Information—The Need for a Policy**

Scientific data may be categorised in many ways: by form or coding (numeric, symbolic, still image, animation, audio, etc.); by content; by means of generation; by level of quality and complexity; by the source of support for the data-accumulating activity; by time and space (observational, geospatial data); and by the institutional structures through which the data are distributed and stored. Certain of these characteristics, such as level of quality (including degree of review and certification) and institutional origin, have given rise to additional challenges associated with the increasingly pervasive electronic communication technologies (NRC 1997).

Some data issues are more discipline specific. Perennial problems affecting access to data include gaps in coverage or quality control, incompatibility of data streams, inadequate documentation of data sets, and difficulty in meeting the requirements for long-term management of data. In the biological sciences, in particular, the variety of attributes and qualifiers included with each observation, the exceptional conditions that need to be catered for, and differences in terminology and usage put a heavy burden on any supplier of data to identify and specify the character of the data precisely enough to prevent misinterpretation. Fragmentation of data into numerous, autonomous, and often incompatible databases, with different formats and levels of quality, is a chronic problem. Putting data to use rapidly in sectors outside the immediate discipline of origin poses additional challenges (NRC 1997).

Only a small fraction of records that have been collected are easily accessible. Many data sets are never archived or exchanged at all. As a result, comparatively few observations are available within an adequate timeframe to support operational needs, respond to emergencies (threatened species, forest fires, oil spills, etc.), or contribute to planning and research management. The quality of data is uncertain and, in most cases, critical metadata are not available. The present users of data are faced with a confusing array of data sets and data formats. The lack of

applied standards results in high data integration and interpretation costs. Overlapping and confused data custodianship and corporate mandates result in further costs from duplication of effort, leading to poor decisions and lost opportunities. Wider understanding of key data management principles and the implementation of a data policy will assist in improving this situation.

A data policy should provide the coordination mechanism to more effectively integrate the activities of the data collectors, data managers and information providers through a distributed data and information network. It should describe an infrastructure and sets of standards, guidelines and procedures to improve the effectiveness of data management and the creation and dissemination of priority data and information in support of corporate goals. Data and information are seen as very valuable strategic assets.

## **1.4 Data Management in Context**

The effective conservation, management and sustainable utilisation of the world's biodiversity resources, for example, are dependent on an adequate data and information base, from which socially, economically and environmentally sound decisions can be made. Data and information are key components in the development of our understanding and the utilisation of these resources within sustainable development principles.

Biodiversity data are collected by a very wide range of organisations and individuals for a multitude of purposes. The majority of data is collected as a result of scientific or educational studies, donor-funded projects, monitoring activities and environmental impact assessments. The cost of data collection is extremely high. However, while this data potentially creates an extensive global data resource, its lack of organisation and coordination reduces our capacity to address a wide variety of user requirements.

In general, the collection, quality assurance, management, interpretation, exchange and dissemination of data is poorly coordinated, inadequately resourced and generally given a low priority. The significance, size and potential value of the data are very substantial, but they need to be, considering the vastness of the earth, with its highly diverse living and non-living resource assets. However, at the present time, integrated data and information systems are so underdeveloped that the available data and information cannot be readily applied to solve our environmental questions at any scale, from local to global.

## 1.5 The Significance of a Data Policy

Effective utilisation and protection of the environment depends on a clear understanding of that environment, ensuring that environmentally sustainable development occurs and 'conflict of use' issues are resolved from a position of knowledge. The development of policy concerning social, economic and environmental issues in this domain is dependent on a comprehensive 'environmental information infrastructure' or 'knowledge base'.

Part of this 'knowledge base' comprises the existing (and future) data and information base. To be of use, this data and information must be well organised and easily accessible, consistent and well documented, so that a multitude of users can easily understand the data they are accessing. The concept behind a data policy is to make this data and information easily accessible and in a useable form to all levels of potential users and uses.

The implementation of a data policy will occur only if it becomes an integral part of an organisation's plans and actions and has appropriate institutional and financial support. To be successful, data management must have a higher profile and be supported through a strong institutional mandate. Success is also dependant on resources being available to implement the necessary infrastructure (building on existing capabilities) and it requires endorsed procedures and guidelines. Success will also require a change in the mind-set towards data management. This can be assisted through the provision of capacity building in and support for data management (see Box 1).

### Box 1

#### Building Capacity for Information Management

A Data Policy will help build capacity in the following ways:

- identify existing strengths and weaknesses in information management capacity;
- identify, document and facilitate sharing of best practice in information management;
- maximise coordination and minimise duplication and overlap in dataset construction;
- identify opportunities for synergy and value adding;
- provide a mechanism for dispute resolution.

BCIS Members, like any other organisation, will only give data management the necessary priority if they see tangible direct benefits outweighing the costs. These benefits of investing in better-managed data include enhanced speed and cost-effectiveness in:

- ingesting standardised data from external sources;
- integrating external with internal data;
- generating more comprehensive, and thus more attractive, information products; and
- mobilising data for others, whether under collaborative arrangements or for some public-good purpose.

Overall, improved data management will raise an agency's profile as a competent professional organisation, thus enhancing business opportunities. Only through establishing an effective data and information infrastructure will the maximum use be made of the Members' resources, which are very limited in comparison to the extent of the environmental challenges facing them.

## **1.6 National and International Implications and Related Activities**

The development of a data policy cannot be undertaken in isolation from other related activities. The policy must incorporate and build around the existing structures that are in place or are being implemented through various projects in other subject domains.

Other initiatives provide a source of valuable data, and interaction with these international programs also provides access to technologies and concepts that can assist improving BCIS data management capabilities. Significant quantities of data are collected in areas of relevance to BCIS by many other agencies. It is therefore important that the BCIS members are able to easily access this data and information to support their corporate and wider objectives. The environment does not observe political boundaries and the capacity of BCIS members to operate at the transnational scale can provide valuable data to address large-scale, transboundary environmental issues. Because of this, international protocols, standards and guidelines for data management must also play a role in a data policy.

A number of international scientific, non-governmental and inter-governmental programs and projects are already underway. BCIS Member agencies are involved in many of these. All the major initiatives resulting from Agenda 21 of the Rio Environmental Conference are being actively supported by BCIS Members.

## 1.7 Commercial Potential

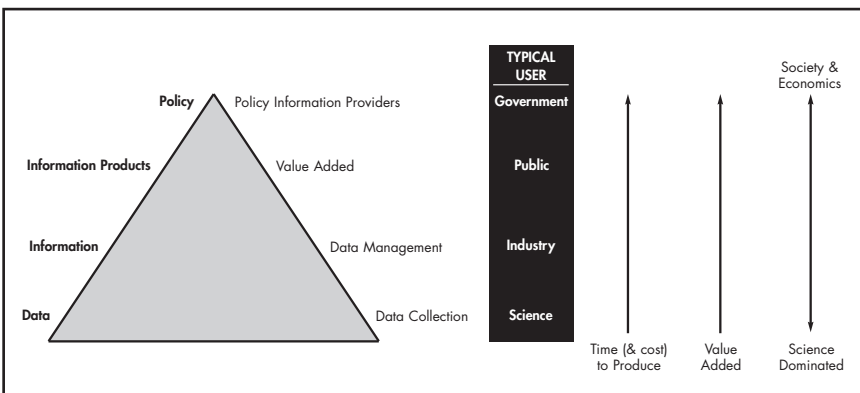
While the focus of a data policy is to improve access to, and utilisation of, data and information, there are other less obvious benefits. The development of a comprehensive data policy, supported by an effective infrastructure, creates a significant number of commercial opportunities for BCIS Members. There are three specific areas that have commercial potential:

- sale of products and services;
- training programs and education (national and international); and
- capacity building through the application of intellectual property and expertise.

## 1.8 The Data—Information Relationship

There is a complex relationship between data and information and the distinction between them is often blurred. Information generation begins with primary data, which are obtained from the environment by various techniques, including instruments, specimen collections and observations. These data are transformed into managed data through processes of quality assurance and incorporation into structured datasets. Data are then integrated, analysed, interpreted and transmitted to users in the form of information. While data is stored and managed in standard form, and may be useful for years, if not centuries, information is very much tailored to the user and has a comparatively short ‘shelf life’.

The following diagram demonstrates some attributes of the relationship between data and information and their respective relevance to a variety of purposes.



*Data and Information—Relationship Diagram*

## 2. GENERAL PRINCIPLES

---

### 2.1 Introduction

A data policy outlines the infrastructure, provides the guidance and establishes procedures and standards for all activities within the data and information cycle. This is achieved using an 'end-to-end' process of management and data flow monitoring (see *BCIS Handbook: Procedures Manual*), from the collection of data, to the management and exchange of data sets, through to the development and dissemination of data and information products. The policy establishes mechanisms for tracking and monitoring data, and provides standards and guidelines to ensure that the maximum utility is made of the data collected by Member agencies and networks. The policy also ensures that the intellectual investment of scientists and data owners is protected.

### 2.2 Goal

The goal of a data policy is to ensure that the process of data collection, data management and data and information dissemination is not only conducted efficiently, in a coordinated manner and in accordance with the Members' objectives, but also results in the capacity to address priority issues at all levels. The policy is achieved through the establishment of infrastructures and procedures based on standards and guidelines developed through a consultative process, while taking into consideration BCIS' wider objectives.

The infrastructure comprises the data collection and data management agencies and their networks, including data custodians and those agencies that process data into value-added datasets and information products. The standards and procedures cover data collection, data processing (including quality assurance), management and dissemination and, importantly, the documentation (metadata) associated with each of these activities. These include data access and licence agreements that specify rights, responsibilities and obligations of various players in the process of transforming raw data into value-added information.

The policy aims to provide the fundamental data and information necessary to provide a balanced view in the conservation and sustainable development debate and the often conflicting requirements of data providers and users.

## Policy Goal

*“An integrated and technologically advanced distributed data management, data and information processing and dissemination capability to support the data and information needs of the organisation, its constituent networks, and decision-makers in developing and implementing environmentally sustainable development policies at all levels.”*

The Policy Goal is achieved through the formalisation of infrastructure, creation of guidelines, standards and procedures to support data management and processing and the generation and distribution of data and information products. This process will be implemented by a well-supported, science-based and technically competent distributed group of data custodians and related agencies. The data management process will be integrated with other networks and initiatives such as the Clearing-house Mechanism of the Convention on Biological Diversity.

## 2.3 Strategies

A data policy will encompass the following strategies to support its Goal:

- create an easily-accessible, distributed data network to manage and disseminate data collected as part of projects and other activities in support of network objectives;
- develop core data sets as baseline products in support of biodiversity conservation;
- provide ‘best practice’ quality assurance mechanisms and procedures to create validated, well-documented data sets to meet priority information needs;
- archive, where appropriate, all data collected, to ensure availability of the data for multiple use and to safeguard it for future generations;
- improve the cost-effectiveness of agency policy and program development through the coordination of data and information activities;
- provide timely and operationally significant data and information products to support a range of activities, with an emphasis on supporting all sectors of society in developing environmentally sustainable development and economic processes;
- make Members’ data available to the widest range of users, at the highest quality and in as timely a fashion as practicable; and
- provide information services and products in the public interest.

## **2.4 The Data Management Process**

Data management is a continuous process, which starts with the design of the measurement, monitoring or data collection program and ends with user access to a comprehensive, quality controlled and fully documented dataset. Data management is a key-component of any scientific or monitoring program. Data management also has a 'value adding' function, which helps to maximise the return on investment by providing the opportunity for multiple use of datasets through effective management and data integration techniques.

Data management is a specialised activity undertaken within a framework of internal and external programs, guidelines, standards and policies.

## 3. PROPOSED STRUCTURE

---

### 3.1 Introduction

The BCIS data-management principles cover the following main elements:

- a coordination mechanism;
- the physical infrastructure of Member agencies and partner organisations that make up the network;
- core datasets;
- analysis and dissemination tools; and
- the standards, guidelines and procedures by which the relevant agencies operate.

#### Basic Concepts

The principles are based on a number of premises including:

- the 'information infrastructure' is a valuable resource;
- data (particularly the core datasets) should be readily accessible; and
- there are benefits, both to Members and wider society, resulting from managing data for multiple use by a wide range of users and for ensuring its availability for future generations.

### 3.2 Co-ordination

#### 3.2.1 Introduction

Data collection, management and dissemination requires coordination to ensure effective and efficient operation. Further, the continued development and refinement of data management and related processes requires a coordination mechanism that recognises and incorporates the views and needs of all stakeholders. The coordination mechanism will build on existing structures and have an appropriate mandate and reporting requirements.

#### 3.2.2 Data Management Group

##### 3.2.2.1 BACKGROUND

The Data Management Group (DMG), which reports to the BCIS Steering Committee, provides the informal mechanism presently in place to enhance data management. It is vital that the DMG

provide an active conduit between management and the information infrastructure and not act as an isolated technical group. The DMG is open to persons from the Member agencies who are willing and able to assist development of the information infrastructure. Persons from non-Member agencies who have the potential to add significant value may also be invited to participate.

#### 3.2.2.2 PROPOSED STRUCTURE

The BCIS Steering Committee, constituted in accordance with the *Biodiversity Conservation Information System: Agreement of Principles*, is responsible for broad policy guidance. The Data Management Group, comprising individual(s) from each Member, coordinates and advises on technical developments. A number of informal technical advisory groups may be used to extend the expertise available to the DMG in specific discipline areas, such as taxonomy, quality assurance procedures, standards, metadata, data exchange formats and communications technologies. The DMG will implement the established priorities to create draft procedures and guidelines and submit these for endorsement by the Steering Committee.

To achieve its objectives of drafting and overseeing the implementation of data management policy, the DMG will use an extensive consultation process with each Member and within the wider research and other relevant communities. This will ensure that the proposals are workable and are consistent with international standards, conventions or best practise. The DMG will oversee the incorporation of relevant standards and protocols into BCIS data policies and guidelines.

A significant role of the DMG is to represent the BCIS in other technical fora, to ensure that, when these groups are creating guidelines and policies, BCIS-Member interests are adequately represented. The DMG will also ensure relevant guidelines already in place are explained and interpreted in appropriate contexts.

### 3.3 Data Management Infrastructure

#### 3.3.1 Introduction

The infrastructure comprises a number of components, including:

- data collection agencies;

- data owners;
- data custodians;
- information agencies; and
- physical infrastructure, including communication linkages such as Internet.

### 3.3.2 Data Collectors

Agencies, including BCIS Members, and individuals collect data for a wide range of purposes. Various multi-disciplinary and (multi-)national data-collecting programs are in place, also making a contribution to the 'Global Information Infrastructure'. Ensuring multiple use of data and simplifying its accessibility to other users begins with the collection process. It is essential that the data collection activities of BCIS Members follow guidelines and procedures in accordance with agreed principles. By promulgating best practice, BCIS hopes to ensure that a level of standardisation and, more specifically, documentation is reached that will reduce future effort and expense, minimise redundancy and maximise opportunities.

It is the responsibility of the collecting agencies to ensure that the data collected during their field projects is managed and documented to the agreed standard(s). Also, in line with these principles, the data must be forwarded to the appropriate custodian within the agreed timeframe.

### 3.3.3 Data Owners

The issues relating to data ownership are complex, especially where the data owners are not the custodians (see *BCIS Handbook: Custodianship*). Data owners have the absolute right, except where qualified by agreement, to decide who has access to their data and under what circumstances. Others hold the view that all environmental data is a fundamental component of the 'global commons' and should be freely available to all. A data policy should recognise this tension and encourage the provision of data into the public domain, recognising that any intellectual property rights or other sensitivities legitimately claimed by the owner must be respected.

### 3.3.4 Data Custodians

#### 3.3.4.1 INTRODUCTION

It is essential that the data management function is built around existing capabilities and expertise. The quantity and diversity of types of data being collected within the domain of interest to BCIS are too great for any single Member to effectively undertake in isolation. The only viable solution is through a distributed system of 'custodians' (see *BCIS Handbook: Custodianship*). Custodians are defined as: *"the nominated body, or person, responsible for the development and/or the management of a dataset, and who has the right to determine the conditions on which those data may be used or released"*.

Custodians are discipline-specific, leading agencies that take charge of data management activities within their core domain. Through implementing the concept of 'custodianship', a number of agencies will be 'endorsed' to manage specific datasets on behalf of BCIS. Endorsement of these agencies will require a process of review and consultation, at management level, through the BCIS Steering Committee [see *BCIS Handbook: Custodianship*].

Custodian agencies could be of any kind, including individuals. However, it will be essential that their role is clearly acknowledged and endorsed and that they are able to 'guarantee' that adequate resources will be available to undertake the task. Data management is a long-term activity requiring on-going resources that must not significantly fluctuate with changes in political or organisational direction.

Custodians require endorsement at the Steering Committee level and will operate under guidelines and standards developed by the Data Management Group, as endorsed by the Steering Committee. The DMG performs an overseeing, coordination and consultative role with the custodians, consistent with data-management principles. The DMG supports the custodians: monitoring the resources allocated to the data management task, ensuring that procedures and guidelines were followed, and overseeing and reporting on the effectiveness of their data management.

The basic infrastructure to develop such a distributed system of custodians is already in place, with a number of agencies being acknowledged as domain experts. These agencies have already created significant infrastructures for the processing and management of their own data. This capability requires only limited expansion to address wider BCIS needs. For instance, a number of

BCIS members already undertake a data custodian role and have documented a number of datasets in the BCIS Metadatabase.

The creation of a BCIS data network will occur through the formalisation of the custodians' wider role under this Policy and with a resource commitment from the custodians. Many agencies already contribute towards meeting wider network obligations, e.g. Botanic Gardens Conservation International, International Species Information System, IUCN–Species Survival Commission, The Nature Conservancy. This process needs to be formalised and strengthened.

#### 3.3.4.2 DATA CUSTODIAN RESPONSIBILITIES AND OBLIGATIONS

The data custodians will be required to meet certain obligations, which include:

- building a dataset (with partners, as appropriate);
- maintaining a dataset (i.e. keeping it up to date, assuring its quality, keeping it abreast of standards and best practise, restructuring it as necessary);
- securing a dataset (i.e. protecting physical integrity from loss or damage, including unauthorised access);
- providing appropriate documentation of the dataset (i.e. metadata);
- providing appropriate access to a dataset (to legitimate users);
- providing advice on appropriate uses of a dataset (e.g. suggested/unwise/improper uses); and
- coordinating further development of a dataset (with appropriate partners and with guidance from key users).

#### 3.3.5 Information Agencies

Many users do not require access to individual datasets, i.e. raw, unprocessed data. Policy makers within governmental agencies, the general public and, in many instances, educational institutions, require processed data and information products. These products come from the integration and analysis of a range of different data types, often across a range of disciplines. This is the 'information manufacture' component of the data and information cycle (World Conservation Monitoring Centre 1998).

A number of agencies already perform this task of 'value adding'. This is achieved through processing and merging data using data fusion techniques to produce information for a wide range of purposes, including policy development, priority setting, decision support and awareness raising. Other agencies undertake significant data fusion processes on a wide range of multidisciplinary datasets to support information requirements at a more political level. The full range of 'information providers' must be considered and their activities incorporated within the BCIS information infrastructure.

### **3.3.6 Physical Infrastructure**

#### **3.3.6.1 ELECTRONIC NETWORKS**

The development of a 'physical' infrastructure is required to bring together the various distributed elements such as the data custodians and information agencies. The most comprehensive and flexible mechanism is using the Internet and its World Wide Web (WWW) capabilities. Most of the existing 'lead' agencies have Internet and database management capabilities, although resources will be required to link these to form a distributed, seamless system. Many of the major clients are also connected to the Internet. However, care must be taken to ensure that technology is not used to the detriment of those without such access. Other mechanisms must also be used to distribute data and information, such as hardcopy publications and CD-ROMs.

It is essential that resources are secured to create this distributed system and that standards and protocols are developed to ensure ease of use and interoperability.

#### **3.3.6.2 TECHNOLOGY**

To implement the above strategies (see 2.3), it will be most cost effective to build on the existing physical infrastructure and use existing expertise in data management and data and information dissemination. Significant advances in computing, particularly distributed computing concepts, available via and supported by the Internet, now make it possible to create an effective electronic network of data managers and information providers.

It will be the Data Management Group's responsibility to ensure that the Internet capability is built on effectively, and that other media such as publications and CD-ROMs are made available to address the full range of dissemination options. The DMG will

also provide the guidelines to facilitate the electronic interconnection between agencies and databases.

Resources will be required to establish this distributed data and information system. However, some of the existing capability is already in place, which means resources can be allocated to creating the links, filling gaps and introducing standardised practices, rather than for large-scale infrastructure development. Expertise already exists within the membership and there is considerable scope for more effective sharing of experience and advice.

The development of this distributed data management infrastructure will also provide significant incentive and experience for the development of other thematic and regional networks, facilitating the development of an interoperable 'network of networks'.

### 3.4 Core Datasets

While the management of all the data collected in areas of interest for BCIS is important, it is necessary to provide a focus and priority (particularly for the allocation of resources) towards specific datasets. These are considered 'fundamental' or 'core' datasets that are necessary to provide a baseline framework for the subsequent development of specialised datasets, information products and services. The composition of a core dataset will need to be determined through a consultative process, which will be facilitated then documented by the DMG [see *BCIS Handbook: Core Datasets*]. Examples of the types of data that are likely to be included as core datasets would include:

- country and other political/administrative boundaries (including protected areas);
- climate;
- basic topography, including catchments;
- surface hydrology (rivers, lakes, etc.);
- substrate (geology, soils, etc.);
- human infrastructure (settlements, population distribution, transport and communication corridors);
- land use, land cover; and
- biodiversity regions (including biomes, ecosystems, vegetation types, etc.).

It is important that these (and other) core data sets are produced from the best available sources and made readily accessible to the full range of potential users through all appropriate mechanisms including the Internet. However, it is important that care is taken to ensure that the spatial and temporal resolution of these fundamental data sets address most user requirements, without compromising data integrity and standards.

### **3.5 Analysis and Dissemination Tools**

Just as in the case of datasets of priority interest for BCIS, it is also necessary to provide a focus and priority (particularly for the allocation of resources) towards specific tools. These would include those that are necessary for basic data analysis through to those required for the development of specialised datasets, information products and services. The key tools will need to be determined through a consultative process, which will be facilitated then documented by the Data Management Group [see *BCIS Handbook: Tools & Technologies*]. Examples of the types of tools that are likely to be required include:

- word processors and spreadsheets;
- database management systems (for both large complex datasets and ad hoc project datasets);
- spatial data management systems (including GIS);
- graphics, video and audio management systems;
- statistical, modelling, expert and decision-support systems; and
- information packaging and dissemination systems (for WWW, CD-ROM, etc.).

### **3.6 Standards, Guidelines and Procedures**

#### **3.6.1 Introduction**

The efficient collation and analysis of raw data and the development of integrated information products requires attention be paid to the following issues:

- quality assurance;
- standards;
- metadata (data documentation); and
- data exchange.

### 3.6.2 Quality Assurance

Data quality is a measure of ‘fitness for use’ and is therefore relative to that intended use. Data can be high quality for one purpose yet low quality for another. While data should be of the highest practicable quality, it is more important to document quality, so a user has sufficient information to decide whether or not those data are fit for their intended use, than to strive for some theoretical quality ideal [see *BCIS Handbook: Standards & Quality Assurance*].

### 3.6.3 Standards

The objective of standards is to ‘reduce the transaction costs of using data’. Again, the objective is to facilitate data integration through data harmonisation, rather than to impose bureaucratic constraints and unnecessary costs on data custodians [see *BCIS Handbook: Standards & Quality Assurance*].

Within BCIS, existing standards will be adopted or adapted. The Data Management Group will look to support development of other appropriate standards, in collaboration with others. Only as a last resort will standards be developed within BCIS, and these are likely to be only interim ones.

### 3.6.4 Metadata

Data documentation is being increasingly recognised as an important component of responsible data management [see *BCIS Handbook: Metadata*]. Metadata contributes to the following objectives:

- marketing information about the datasets held by a custodian or network;
- assisting internal management of datasets; and
- providing detailed descriptions of data elements to support user applications.

### 3.6.5 Data Exchange

These principles strongly recommend unrestricted and ready access to the core data sets by all legitimate users, in the wider public interest [see *BCIS Handbook: Data Access*]. Data exchange should be managed through data exchange agreements. The purpose of these agreements is to expedite rather than inhibit data flow. Issues relating to intellectual property and cost recovery need to be addressed.

Whether or not to charge for data and, if so, what pricing policy should be used is a complex issue. In general, the secondary user of data should not pay the collection cost of the data a second time (double counting). Maintenance costs borne by a data manager are an overhead cost of doing business and should be recovered, like other overheads, from the agency's overall revenue stream, rather than directly from data users [see *BCIS Handbook: Data Access*].

It is an objective of the data-management principles that the maximum return is made of the expenditure outlaid on the collection and analysis of data, through the multiple use of those data. The ability to leverage from the initial investment through effective data management is a focus of the BCIS information management framework. While this may incur an additional cost, the benefits are considerable when avoidance of duplication and the cost of lost opportunities are considered. The significance of opportunity costs should not be underestimated when looking at environmental decisions being made using inadequate data and information.

One of the most significant obstacles to data exchange is communication difficulties between custodians and users—understanding the issues and negotiating equitable data access agreements [see *BCIS Handbook: Data Access*]. Another important obstacle is the diversity of data formats used by collection and data management agencies. The main challenges are:

- the measurement of different environmental entities in the same place, or of the same entities but in different ways, for different (often poorly documented) purposes;
- differing ranges or allowable values for attributes of those entities;
- incompatible database structures; and
- different data exchange formats.

A major responsibility of the Data Management Group is to assist in the development of standards for data exchange, so that valuable resources are not wasted on converting data from one system to another. Unfortunately this is a very complex issue which will not be solved quickly. However, the international data management community is attempting to address this problem and BCIS should continue its monitoring of and cooperation with relevant activities.

The developing use of the Internet and the concept of distributed databases reduces aspects of this format problem but does not solve it entirely. The Data Management Group will be required to focus on developing a solution to the various dimensions of the data exchange issue.

## 4. IMPLEMENTATION CHALLENGES

---

### 4.1 Freeing up the Flow of Data

One of the barriers to the implementation of a data policy is a general reluctance of data collectors to submit data to a 'custodian'. This occurs for many reasons, but the primary factors are:

- belief that the data 'belongs' to the collector and that only the collector really 'understands' it; and
- few incentives to make the data available for other users.

One mechanism to overcome this comes through the funding sources. Most projects require a proposal for funding that is often reviewed externally. A data policy might suggest that funding program guidelines include a section on data management activities, describing the process that will be followed and indicating to which custodian(s) the data will finally be submitted. This data management plan should also identify resources required to create a data set of adequate quality (addressing appropriate standards and including metadata) to be lodged with a custodian. Management approval for project proposals could be made dependant on following this procedure. By linking provisions for long-term data management, in this instance lodgement of data, to the budget process, project staff are able to identify resources required to perform this task and also accept the responsibility for submitting data.

The BCIS data-management principles recommend that a data management plan forms a part of every project [see *BCIS Handbook: Procedures Manual*].

Where data has further operational uses, the project's data management plan must identify the mechanism by which this data are made available to the appropriate operational agency in near real-time. The final data set must be forwarded to the custodian agency within a prescribed time, e.g. within 2 years. This covers the 'right to publish' issue, providing sufficient time for analysis and publication while meeting wider needs.

The designated custodian(s) must undertake monitoring of this data flow to ensure that all the data collected are appropriately archived. The Data Management Group will coordinate the task and assist the custodians with the monitoring process.

## 4.2 Resources for Data Management

Finding the resources for current levels of data management presents challenges to all organisations, including BCIS Members. Finding the staff and financial resources necessary to implement a comprehensive data policy and participate effectively in a collaborative network can be very difficult. Even though most managers accept, in principle, that major efficiency and effectiveness gains can be made through better organisation and collaboration, these gains are in the medium- to long-term, while the costs are incurred 'up front'. Seed funding from external sources, such as donors or program funds, can be crucial in expediting the process. Where these are not available, progress can be disappointingly slow.

## 4.3 Resolving Ownership and Custodianship

Organisations involved with data management often have a long history of working with others in a variety of bilateral and wider arrangements. Rarely are these arrangements on a strategic, long-term footing. Mostly they were initiated and have evolved haphazardly, according to the imperatives at the time. Many of these arrangements are fragile and marked by tension, occasionally conflict. Even where formal written agreements are in place, which is seldom, these rarely address all the salient points and often the language used confuses or inhibits the resolution of outstanding data management issues.

We now face a daunting legacy of datasets that have been contributed to by a number of parties, but where the ownership and custodianship issues have never been addressed, let alone resolved. Unless sensitively managed, these datasets can be a source of friction among the parties.

There are two levels at which these situations can be addressed. At the operational level, mutual understanding of the key principles of data ownership and custodianship [see *BCIS Handbook: Custodianship*] can assist. Ultimately, however, these issues must be agreed, if not resolved, at management level and subsequently endorsed by the Steering Committee.

## FURTHER READING

---

Australian Marine Industries and Sciences Council–Marine Industry Development Strategy.

Babbitt, B. 1995. Science: Opening the Next Chapter of Conservation History. *Science* 267: 1954–5.

Commonwealth Spatial Data Committee Annual Report 1995-1996, Australia.

IGOSS/IODE Data Management for the Global Ocean Observing System.

NRC, 1997. *Bits of Power: Issues in Global Access to Scientific Data*, National Research Council, USA, 1997.

Preliminary Report of the Information Management Steering Committee on Information Management in the Commonwealth Government, Australia.

State of the Marine Environment Report, Australia.

World Conservation Monitoring Centre, 1998. *WCMC Handbooks on Biodiversity Information Management* (8 volumes). Reynolds, J.H. (Series Editor). Commonwealth Secretariat, London.