

BIODIVERSITY DATA MANAGEMENT
(Document 2)

GUIDELINES
for a
NATIONAL INSTITUTIONAL SURVEY

in the context of the
Convention on Biological Diversity



**WORLD CONSERVATION
MONITORING CENTRE**



**The mission of the
World Conservation Monitoring Centre is to provide
information on the status, security and
management of the Earth's biological diversity.**

9878

BIODIVERSITY DATA MANAGEMENT
(Document 2)

GUIDELINES
for a
NATIONAL INSTITUTIONAL SURVEY

in the context of the
Convention on Biological Diversity

United Nations Environment Programme

March 1995

Digitized by the Internet Archive
in 2010 with funding from
UNEP-WCMC, Cambridge

ACKNOWLEDGEMENTS

1. This document is one of a series of four researched and compiled by the World Conservation Monitoring Centre, Cambridge UK with 80% funding from the Global Environment Facility (GEF) through the United Nations Environment Programme (UNEP), Project GF/0301-94-40 (GF/0301-94-06). The need for the development of a package of tools and materials to support national information management for the Convention was identified and the project promulgated by Mark Collins (Director, WCMC) and Robin Pellew (former Director of WCMC).
2. Principal authors were Donald Gordon and Chris Magin. Contributions and critical review were received from a number of WCMC staff and consultants, including Christine Carey, Ian Crain, Gareth Lloyd, Gwynneth Martin, and Jake Reynolds. The document has benefited, as well, from review and comment from NGOs, UNEP, and experts in a number of countries including Colin Bibby of Birdlife International, and others who participated in a consultation meeting hosted by UNEP in Nairobi in October, 1994. Graphical concepts were developed by Gwynneth Martin, Ian Crain and Don Gordon and executed by Ian Kime of "Constructive Solutions". Document organisation, integration and input was by Laura Battlebury. Ian Crain was the project manager and responsible for overall design and editing.

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Background	1
1.2	Purpose of the Institutional Survey Guidelines	2
1.3	The Context	3
1.4	Basic Principles of the Institutional Survey	4
2	CONDUCT OF THE SURVEY	7
2.1	Approach	7
2.2	Lead Responsibility	7
2.3	Steering Committee/Project Team	7
2.4	Preliminary National Workshop	10
2.5	Information Collection	11
2.5.1	Use of Questionnaires	11
2.5.2	Questionnaire Design and Content	12
2.5.3	Questionnaire Development	13
2.5.4	Questionnaire Implementation	13
2.5.5	Conduct of Site Visits	14
2.6	Information Storage and Retrieval	16
2.7	Analysis and Reporting	16
2.8	Finalising the Process	17
2.8.1	Final Workshop	17
2.8.2	Distribution of Final Report and Associated Products	18
2.9	Follow-up	18
3	DEFINING THE SCOPE	19
3.1	Institutions to Include	19
3.2	Typical Institutions to Survey	20
4	CATALOGUING BIODIVERSITY INFORMATION HOLDINGS	23
4.1	Approach	23
4.2	Compilation and Presentation of Data	23
4.2.1	Production of Directory	23
4.2.2	Directory Contents	23
4.3	Metadatabase Construction	24
4.3.1	Metadatabase Contents	24
4.3.2	Metadatabase Formats	25
4.3.3	Metadatabase Data Entry	25
4.3.4	Metadatabase Updating	26
4.3.5	Metadatabase Documentation	26
5	ASSESSING INSTITUTIONAL LINKAGES	27
5.1	Approach	27
5.2	Linkage Information Collection	27
5.3	Information Compilation and Presentation	29
5.3.1	Institutional Level	29
5.3.2	National Level	30

6	ASSESSING INFORMATION MANAGEMENT CAPABILITIES	33
6.1	Approach	33
6.2	Data Collection	33
6.2.1	Human Resources	33
6.2.1	Technical Resources	33
6.2.3	Financial Resources	34
6.3	Information Compilation and Presentation	35
6.3.1	Institutional Overview	35
6.3.2	National Overview	36
6.4	Assessment	36
7	REFERENCES	38

LIST OF ANNEXES

1:	SAMPLE DIRECTORY ENTRY	39
2:	MODEL QUESTIONNAIRE: NATIONAL INSTITUTIONAL SURVEY	41
3:	MATRIX FOR IDENTIFYING INSTITUTIONS MANAGING BIODIVERSITY	53
4:	MODEL TABLE OF CONTENTS: INSTITUTIONAL REPORT	55
5:	MODEL TABLE OF CONTENTS: FINAL NATIONAL REPORT	57
6:	TARGET INSTITUTIONS IDENTIFIED IN EAST AFRICA	59
7:	MODEL METADATABASE FORMAT	61
8:	LIST OF ACRONYMS & ABBREVIATIONS	77

1 INTRODUCTION

1.1 Background

The *Convention on Biological Diversity* (CBD), signed at the Earth Summit in Rio de Janeiro in June 1992, is a landmark in international conservation. Article 6 of the Convention requires ratifying countries to prepare national strategies, plans, or programmes for the conservation and rational use of their resources. An important first element in this process is the preparation of a *Country Study*, whose components are as follows:

- identification of components of biological diversity that are important for its conservation and sustainable use
- collection and evaluation of data needed for the effective monitoring of the components of biological diversity
- identification of processes and activities that threaten biological diversity
- evaluation of the potential economic implications of the conservation and sustainable use of biological resources
- determination of the economic values of biological and genetic resources
- assessment of priority actions for the conservation and sustainable use of biological diversity.

In 1993, the United Nations Environment Programme (UNEP) with the assistance of the World Conservation Monitoring Centre (WCMC) prepared the *Guidelines for Country Studies on Biological Diversity* (UNEP, 1993), intended to assist countries in assessing the status and value of their biodiversity and biological resources, and in developing strategies and action plans to safeguard them. The document provided guiding principles, followed by recommendations on information collection and management presented in a technical annex, which was divided into five sections. Section D, Current Capacity, emphasised that the national assessment must include an overview of the current capacity within the country for the conservation and sustainable use of biodiversity. It emphasised that particular attention should be given to the following:

- institutional and organisational capacities
- human resources
- national environmental legislation
- technological facilities
- information resources.

The effective planning and implementation of biodiversity protection programmes is therefore supported through identification of **sources and types** of biodiversity information available for a country, as well as through assessment and building of institutional capacity.

Having recognised the need for management of data in support of national planning related to biodiversity, the United Nations Development Programme (UNEP), in collaboration with the World Conservation Monitoring Centre (WCMC), designed and submitted to the Global Environment Facility (GEF), a project proposal entitled *Biodiversity Data Management Capacitation in Developing Countries and Networking Biodiversity Information (BDM)*. This proposal was endorsed and subsequently a sub-project was established between UNEP and WCMC for Development of Supporting Materials for Biodiversity Data Management and Exchange.

The sub-project has produced an interlinked package of resource supporting materials to assist in national capacity building. There are four principal components of this package:

Document 1. Data Flow Model

- to identify in a formal structure the relationships between components of biodiversity data, from acquisition through to use in national strategy development, planning, and monitoring for implementation of the CBD.

**Document 2. Guidelines for a National Institutional Survey
(This Document)**

Document 3. Guidelines for Information Management

- to facilitate the development of capacity for information management and exchange as required by the Convention.

Document 4. Resource Inventory

- the core output of the project; a collection of reference directories, guidelines, and standards relating to biodiversity information management.

1.2 Purpose of the Institutional Survey Guidelines

These guidelines are intended to assist countries which are Parties to the CBD to assess the current state of their institutional capacity for managing biodiversity information. In many respects these guidelines are an extension and elaboration of the summary guidelines on determining institutional capacity given in Section D of the *Guidelines for Country Studies*, and, where a Country Study has been completed, are meant to build on this base, and on recently completed initiatives, such as the study entitled *Availability of Biodiversity Information for East Africa* (WCMC, 1993). The Institutional Survey will identify a broad range of institutions within the country that are involved in the collection and management of biodiversity information. It will explain each institution's involvement, and will describe existing information management procedures, including hardware and software used. The guidelines outline how this survey should be conducted and information collected and organised. It is nevertheless recognised that each country has its own particular circumstances and requirements. These guidelines are not therefore intended as an inflexible set of rules and regulations, but as guiding principles which participating countries may elect to follow.

1.3 The Context

The objectives of the Convention on Biological Diversity, given in Article 1 are:

"the conservation of biological diversity, the sustainable use of its components, and the fair and equitable sharing of the benefits arising out of the utilisation of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding".

Sustainable use of biological resources can only be achieved by a sound system of strategic planning. Information must be gathered and analysed; strategic objectives must be set on the basis of this analysed information; action plans and targets should be developed; and concrete actions should be introduced to implement them. In turn, information on the effects of these actions should be gathered and analysed, and "fed back" into the process to allow the evaluation of performance and the modification (if necessary) of strategic objectives and action plans details. This process has been illustrated in diagrammatic form in the *Data Flow Model* (Document 1) in Figure 1.1. The institutional capacity of a country is a vital element in this process, determining its ability to accomplish the planning process efficiently.

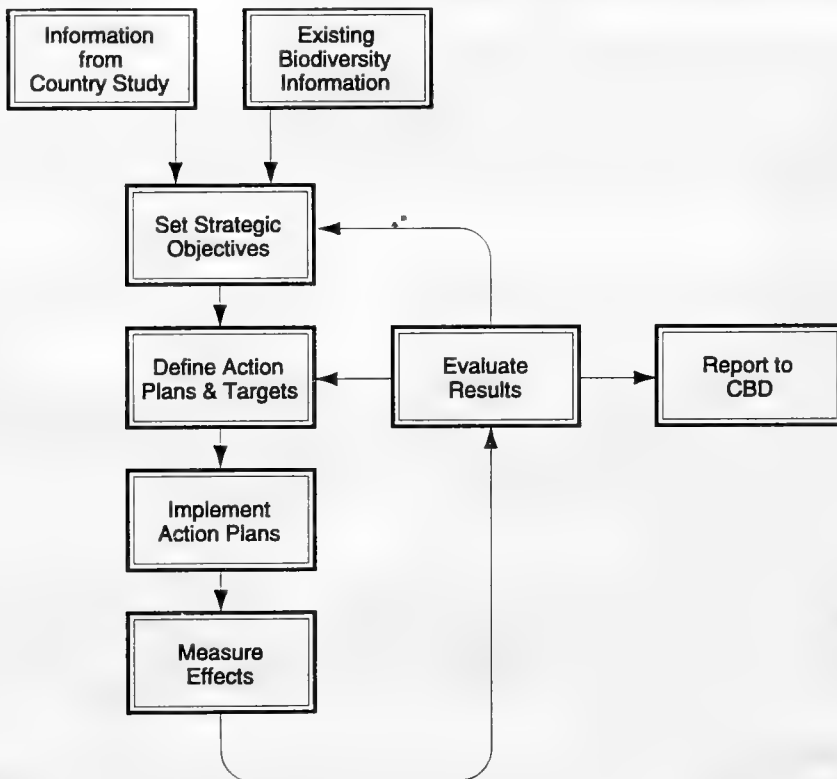


Figure 1.1: Overview of National Activities in Support of the CBD

The *Data Flow Model* establishes, in conceptual terms, the types of information needed and the relationships between institutions involved in biodiversity information management in support of the CBD. The Institutional Survey Guidelines represent an important first step in identifying resources, in-country capacity and associated gaps in the context of this model.

1.4 Basic Principles of the Institutional Survey

A number of difficulties are inherent in surveys of any type. In order to avoid some of the more common pitfalls, it is recommended that the national institutional survey should conform to the following principles:

Clear Mechanism

It is essential that the mechanism by which the institutional survey is carried out and its products discussed and prepared is clearly defined at the outset. The mechanism may vary on a country-by-country basis, but a possible procedure is outlined in Section 2.1.

Independence

While all surveys strive to ensure that they are as independent and impartial as possible, the introduction of bias is inevitable. To ensure balance, governmental, non-governmental and private agency personnel should be included in the survey process (see Section 2.3).

Multi-disciplinary and Multi-participatory

Since biodiversity information management encompasses many different subject areas, the survey should adopt a multi-disciplinary, multi-participatory approach, perhaps by inviting experts from several disciplines to constitute a Steering Committee, which could then advise on the methods to be employed by a smaller Project Team (see Section 2.3).

Cooperation

Collaboration between institutions should be an integral part of the process at all stages of the survey.

Quality Assurance

Care should be taken to ensure that information gathered is reliable. Secondary sources should be avoided. Data gathered by indirect methods, (eg posted questionnaire) should be validated by direct methods (eg interviews during site visits).

Repeatability/Monitoring

The information collected in the survey will act as a benchmark against which institutional capacity development can be measured. In addition, it is hoped that on completion of the survey, a mechanism will have been set in place which will allow periodic or continuous updating of information on institutional capacity.

Compatibility

The survey should be used wherever possible to promote the use of data standards which will facilitate the exchange of data between both national and external agencies (eg regional/international centres).

Facilitation

The products of the institutional survey should facilitate future networking between national institutions and external agencies, and promote increased exchange of datasets.

2 CONDUCT OF THE SURVEY

2.1 Approach

The conduct of an institutional survey will be highly dependent on circumstances in each country. A flow diagram illustrating a suggested process is given in Figure 2.1, and further details of the steps involved are given in the following sections. Elements of this process can be incorporated into the actual mechanism decided on by the host country. The overriding goal of any process adopted, however, is the production of an output which will be of benefit in assisting countries to meet their obligations as set out in the CBD.

2.2 Lead Responsibility

Although the institutional survey is a cooperative venture involving many institutions, it is important that a single recognisable unit is designated as taking lead responsibility for its implementation. UNEP's *Guidelines for Country Studies* (UNEP, 1993) recommend that each country undertaking a country study should establish a multi-disciplinary National Biodiversity Unit (NBU) to co-ordinate biodiversity management, utilisation and research. A number of countries have already completed the country-study process and have created NBUs. If such an institution exists in-country, it would be an obvious choice to take the lead for conducting the national institutional survey. In the absence of an NBU, an existing government department, (eg a Wildlife or Game Department, Forest Service or Environmental Agency) could be considered.

"Who" is chosen is, nevertheless, secondary to "how" an institution is identified as the lead agency. Great care needs to be taken to ensure that institutional perspectives and rivalries do not bias the survey. Options for arriving at this decision include:

- appointment of the lead organisation by an institution charged with overall responsibility for biodiversity matters
- round table discussions based on strengths, weaknesses and capacity of potential lead institutions; choice then arrived at by vote
- informal discussions between institutions during the preliminary national workshop or other appropriate forum, arriving at a lead institution by consensus.

Guideline 1

Identify an in-country organisation to take lead responsibility for conducting the institutional survey.

2.3 Steering Committee/Project Team

The first action of the lead institution should be to establish a small Steering Committee of about 5-10 individuals. It is unlikely that any single department or institution will contain a wide enough range of expertise to cover all disciplines to be surveyed. The need for impartiality and independence, coupled with that for a broad-based, multi-participatory approach, would suggest that such individuals should be selected from a variety of disciplines and organisations, so that they are representative of the broad community of biodiversity information managers and users. It is suggested that an effort be made to include

governmental, non-governmental, and private sector personnel from national universities, museums, and consulting agencies.

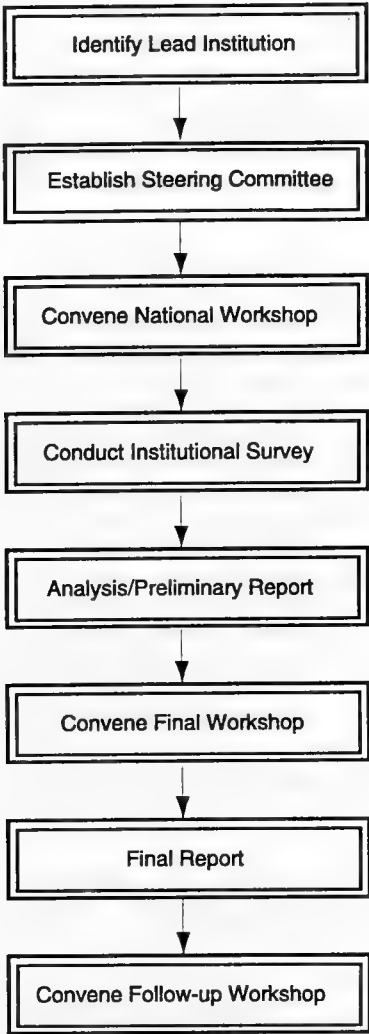


Figure 2.1: Major Steps in Institutional Survey

The Steering Committee would be primarily responsible for the following activities:

- defining the goals, outputs and scope of the institutional survey (see Section 3)
- preliminary planning of survey process and timetable (this section), including discussion of data collection methods (Section 2.5)
- appointing a Project Team (suggested 2-3 people, who may be external consultants) to be concerned with the day-to-day administration of the survey,

- actually carrying it out, analysing the data collected, and preparing outputs
- preparing a list of experts and institutional representatives to invite to a preliminary national workshop to shape the survey process (see Section 2.4)
- supervising the conduct of the survey
- convening a final workshop
- ensuring that appropriate follow up mechanisms are in place.

Model terms of reference for the Steering Committee and Project Team are provided in the boxes below:

**STEERING COMMITTEE
MODEL TERMS OF REFERENCE**

Goals

- Oversee, monitor and advise on project development and implementation.
- Communicate purpose and progress of the survey.
- Develop further initiatives stemming from the institutional survey in support of the CBD.
- Ensure continuity of the initiative by establishing a mechanism (eg schedule and funding) for updating project outputs, supporting capacity building activities, and encouraging networking and accessibility to datasets.

Composition

- 5 to 10 individuals, representative of:
 - the main user communities
 - biodiversity information managers in the governmental, non-governmental and private sectors
 - wide spectrum of institutions that manage biodiversity information in conservation, land-use, legislative and socio-economic areas of expertise.

Activities

- Set project focus.
- Establish timetable.
- Appoint project team.
- Identify institutions to survey.
- Monitor project implementation.
- Convene workshops.
- Initiate follow-up.

PROJECT TEAM MODEL TERMS OF REFERENCE

Goals

- Under the guidance of the Steering Committee, conduct a national institutional survey.
- Advise Steering Committee on the development and implementation of the project, and in proposing follow-up initiatives.

Composition

- 2 to 3 individuals who have:
 - experience in project design, administration, planning and management
 - awareness of the institutions and culture of the country
 - a variety of backgrounds and expertise in biodiversity information management
 - overall grasp of the purpose and orientation of the project in support of the *Convention on Biological Diversity*.

Activities

- Collect information from surveys, questionnaires and site visits.
- Assist in the organisation and planning of national workshops.
- Draft interim and final project reports.
- Analyse data collected with emphasis on taxonomic and geographic coverage of biodiversity information, and assessment of institutional capacity.
- Prepare a catalog of survey results and, and if appropriate, a metadatabase.
- Distribute final project outputs to participating organisations and other interested parties.

Guideline 2

Select a representative group of officials to form a Steering Committee to coordinate and direct the conduct of the institutional survey. The Steering Committee would be responsible for appointing a Project Team to actually carry out the survey.

2.4 Preliminary National Workshop

An important initial task of the Steering Committee is the organisation of a preliminary national workshop of biodiversity experts, users, and representatives of the main institutions which are likely to be surveyed. To ensure the success of the survey, it is vital to establish a spirit of cooperation and teamwork at an early stage. Accordingly, the aims of the workshop are to:

- provide a forum for discussion

- further define the objectives, scope, methodology and activities of the project (already outlined by the Steering Committee)
- discuss outputs and potential benefits
- consolidate the project team, and assign responsibilities to various institutions
- further develop a timetable
- discuss aspects of finance/incentives for carrying out the survey.

Two products that need to be finalised at this stage are a pilot questionnaire (see Section 2.5.3) and a list of institutions to be included in the survey (see Section 3).

Guideline 3

Invite a number of selected biodiversity experts, conservation decision makers, and users from government departments, non-government institutions and private agencies to participate in a national workshop to plan the institutional survey.

2.5 Information Collection

2.5.1 Use of Questionnaires

The approach taken to gathering data is entirely up to the country concerned. However, it is suggested that standard questionnaires and "check-lists" be used either as data collection tools in themselves, or as guides to facilitate data gathering. A properly-designed questionnaire is a valuable tool which promotes the systematic collection, cataloguing and evaluation of data. In the context of the institutional survey, a three-pronged questionnaire can be used to assess the biodiversity information holdings of institutions, ie their datasets (see Section 4), institutional links (Section 5), and information management capabilities (Section 6). The questionnaire approach to collecting data on institutional datasets has recently been successfully tried and tested by a WCMC/IUCN (World Conservation Union) East African Regional Office project entitled *Availability of Biodiversity Information for East Africa* (WCMC, 1993). The model questionnaire presented in this document is heavily based on that experience. An example of the information collected from an institution in the East African survey is given in Annex 1. The design, content, and implementation of a standardised questionnaire to facilitate the collection of data concerning institutional capacity to manage biodiversity information is expanded upon in the following sections. It must be stressed, however, that the use of questionnaires *per se* is but *one* method of obtaining required information.

The use of questionnaires in the form of a simple mailing has proven in the past to have weak response and presented difficulties in interpretation. The principal value lies in having the questionnaire as a data collection form to standardise and consolidate the information gathered by a range of means, or as preliminary screening to help determine which institutions should be studied in more depth.

Guideline 4

Develop standard questionnaires or data collection forms to assist the process of information collection augmented with other methods.

2.5.2 Questionnaire Design and Content

A model questionnaire illustrating the type of information that should be sought from institutions is included in Annex 2. It is suggested that the questionnaire be split into two parts: the first dealing with information pertaining to the institution itself, the second with the datasets it maintains and uses. A suggested breakdown of contents follows:

Part 1: Institutional Information

- Institutional details (eg title, government or non-governmental organisation (NGO), postal address, telephone/fax/telex, email, and brief description of institutional function).
- Institutional resources (eg availability of trained personnel, presence of technical equipment, hardware, software).
- Institutional links (eg source of data, users, how information is exchanged with other institutions).

Part 2: Dataset Information

- Description of dataset (eg name or subject area of dataset, name of information manager, form and size of dataset, reason why it is held, how the information is managed, details of computer systems used including hardware, operating system and software).
- Information coverage (eg geographical area of coverage, biomes, ecosystems etc covered, description of major taxa included, time-span of coverage, completeness, limitations and gaps in coverage).
- Information access (eg availability of information to outsiders, cost of access, formats available, electronic communication, availability of documentation).

A dataset can be defined as a collection of data on biodiversity and its accompanying documentation. Biodiversity should be considered in its broadest sense so that although information pertaining to biology (eg animals, plants and habitats) is of primary importance, information concerning land use, physical features, legal and economic aspects should also be included. For the purposes of this survey, a "dataset" could be a single file, but more commonly will consist of a collection of thematically related files which can be conveniently described as a group. An example is "Forest ecology data from the XX region". It is **not** intended that each individual record, map, report, coverage (GIS) or small electronic file be separately listed and described.

Any questionnaire used for gathering biodiversity information should be as brief as is practicable. Wherever possible, check boxes should be positioned beside answers to questions so that respondents simply need to tick a response, or tables should be used so that answers can be compiled systematically. When institutions hold more than one relevant dataset, a separate Part 2 of the questionnaire should be filled out for each (Part 1 could be left blank

on these "duplicate" questionnaires since the institutional information will be identical for each dataset). Where institutions have more than one department or unit dealing with biodiversity information management, it is suggested that each is treated as a separate entity, and separate questionnaires are completed for each.

At the discretion of the host country, a *Needs Assessment* for capacity building could be carried out at the same time as the *Institutional Survey*. Questions could be added to the questionnaire targeting the institutional requirements perceived as necessary to perform their designated biodiversity information management role. Particular attention could be paid to:

- goal definition
- data needs
- information management needs, (eg technology, systems plan, networking and accessibility)
- staffing and training needs.

Guideline 5

Questionnaire design should be kept simple, and the number of questions kept to a minimum. The questionnaire provided in Annex 2 could be used as a draft model. Separate questionnaires should be completed for each distinct dataset held by an institution.

2.5.3 Questionnaire Development

The Steering Committee should discuss and (as necessary) modify the contents of the institutional survey questionnaire to suit the national context. This resulting draft questionnaire may be further refined following consultation with participants of the National Workshop. The resulting pilot questionnaire could then be tested on a trial basis to a limited number of key institutions - perhaps two or three - and further modified in response to any difficulties encountered before full implementation of the study commences.

Guideline 6

The Steering Committee should present an initial draft questionnaire to the National Workshop, and incorporate its comments and suggested changes.

Guideline 7

The Project Team should implement the pilot questionnaire with a small number of institutions on a trial basis before commencing the full study.

2.5.4 Questionnaire Implementation

One of the aims of the Steering Committee and the National Workshop is to provide an exhaustive list of "target" institutions, ie those involved in the use, management and dissemination of biodiversity information. Guidelines on how to identify target institutions are given in *Defining the Scope*, Section 3. After finalisation, and if it is agreed that a questionnaire should be used for survey purposes, it would then be sent to these institutions. Provision of the questionnaire could be in hard copy or as a digital file. Further relevant institutions will probably be uncovered during the course of the survey, and could be added to the list.

Normally, very low response rates are obtained from blind mailings alone, but these can be increased by including a supporting brochure providing a summary explanation of the purpose of the study and questionnaire, together with a sample questionnaire completed as an illustration. However, even if completed questionnaires are returned by post, respondents may have difficulty answering some of the questions (eg details of the size of their data holdings), left some blank, or filled in the questionnaire incorrectly.

The most effective way of improving the quality of the data obtained is by conducting follow-up "site visits" to the institutions concerned. In the framework of a site visit, the questionnaire can act as an "aide-mémoire" to the interviewer, ensuring that all the relevant questions are asked and enabling the interview to progress logically. It may or may not be possible to visit all institutions identified as holding or managing biodiversity information, depending primarily on how many are involved. If a large number are identified, the Steering Committee could select a sub-set based on the information in the preliminary returns of the questionnaire, and direct the activities of the Project Team accordingly so that site visits are focused on key institutions only.

Guideline 8

Where it is deemed appropriate that a questionnaire per se be used to solicit information, then the Project Team should implement the finalised version. The initial step is to distribute questionnaires by post to target institutions, including a completed example. Questionnaire completion should be facilitated by conducting site visits.

2.5.5 Conduct of Site Visits

Site visits can be prepared in advance via contact with the director/manager of the institution. An initial letter should be sent to reinforce the purpose of the institutional survey and to clearly outline prospective benefits to the host. These include improved accessibility to data resources, financial incentives, and involvement in follow-up initiatives; the point being that unless institutions see the mutual benefit of being involved in such an exercise, the survey will accomplish far less than intended. Some institutions may be sceptical as to the intent of the institutional survey, and may fear loss of control over their data. It should be clearly explained, both in the supporting brochure accompanying the questionnaire and during the contacts, that raw data is not being collected, but rather data about data holdings (ie metadata) and information on institutional strengths in order to *improve cooperation* and enhance the information management capabilities of institutions through an integrated network.

The initial contact should be followed by communications to identify the appropriate persons (the managers of datasets) in the institution, and agree a mutually convenient time and date for a site visit by the survey team. The conduct of the site visit will vary according to the customs and cultures of the host nation, plus the organisational structure and internal hierarchy of the institution concerned. Useful techniques include:

Individual interviews

After identifying the biodiversity information manager in an institution or other relevant person, the questionnaire is then completed in either: (a) a formal sense in which questions are asked and responses recorded on tape or written on the form itself; or (b) in a more

informal sense where the questionnaire is used as a guide to discussing key topic areas. Information can either be recorded at the time or following the interview; issues to consider include accuracy in recording information, and in using an approach which is both culturally acceptable and solicits the maximum amount of information.

Group interviews

A small number of key individuals in an institution who are involved in biodiversity information management are interviewed together. The survey questionnaire is used as a guide to solicit and record information. Information from the group is then summarised in completing an institutional entry for the study. It is useful to have one person to lead the discussion and another to record important information. Individual and small group interviews can be handled casually by talking to various staff members during the course of a whole institutional visit.

Small group meetings

Similar to group interviews, but often involving a larger number of individuals and accommodating a wider spectrum of interests and expertise. The meetings often involve a short presentation on the topic before opening up the discussion more fully. The questionnaire is typically used as a guide to introducing and discussing key topic areas.

Brainstorming session

Effectively used to pull in any and all ideas on a subject from a wide range of interests and expertise. A facilitator is needed to initiate and steer the process, as well as create an atmosphere which stimulates creative thought. All inputs are recorded, information being sorted/used where applicable/relevant in the context of the project. Brainstorming is particularly useful when defining the scope of a project, in support of a directional change to the initiative, or for an infusion of new ideas and inspiration. For example, brainstorming may be useful in trying to identify biodiversity datasets of an institution.

Workshops

Workshops typically vary between one day and one week in duration, and are used to discuss, plan and assess many aspects of a project. These aspects range from methodologies, outputs and progress to follow-up initiatives. Often separate task groups are formed to pursue selected issues. Workshops often incorporate elements of training, and where a wide spectrum of institutions is involved, facilitate sharing of knowledge and expertise. Workshops are useful for arriving at decisions from consensus-building exercises; for example, this would be an effective way of identifying data management processes at the national level.

Enough time should be allowed for thorough visits, particularly where remote institutions may be involved. Due to the potential time and expense involved with site visits, the alternative of holding national/regional/local workshops to solicit survey information should be considered. The project team, with the assistance of staff from key institutions would be able to discuss the survey with participants using any number of the approaches suggested above. The added advantage of this approach is that it would promote active networking between institutions.

Guideline 9

The Project Team should plan and prepare site visits thoroughly in advance, and carefully explain the objectives of the survey to ensure cooperation. The standard questionnaire (see previous sections) can be used as an "aide-mémoire" to obtain the necessary institutional and dataset information. National, regional or local workshops could be considered as an alternative to, or in combination with, site visits.

2.6 Information Storage and Retrieval

For the institutional survey to have lasting analytic value, the information compiled should be made accessible in hard copy form as a "directory." It is also useful to maintain the information electronically in database format. A database containing information about datasets is commonly referred to as a "metadatabase". Any metadatabase developed should be able to accommodate the information obtained via the survey approach, be this questionnaire or otherwise. The software used should allow data to be organised, searched and analysed, for example to identify gaps in biodiversity information available for the country. For each dataset entry, the use of "keywords" is important to enable a compatible search of electronic records in the metadatabase. Standard keywords are defined in such sources as the *INFOTERRA Thesaurus of Environmental Terms* (UNEP, 1990). Metadatabase formats should conform to international standards to facilitate the exchange of data within and between countries. Details of relevant metadatabase/directory formats are given in Cataloguing Biodiversity Information Holdings, Section 4.

Guideline 10

The Project Team should present their findings as a directory, and if possible, also as a metadatabase.

2.7 Analysis and Reporting

Once the data have been compiled, they can be analysed to identify existing strengths and gaps in institutional capacity and the availability of biodiversity data. Suitable themes include:

- Institutional Information Management Capability
 - levels of technical resources
 - levels of human resources.
- Institutional Links
 - links between national/international institutions receiving and supplying data or collaborating (see Section 5.3).
- National Dataset Coverage
 - which institutions hold data
 - extent of coverage
 - details of collection and maintenance of such information
 - taxonomic breakdown of information held (eg fish, plants)
 - coverage by biome (eg freshwater/marine/desert/tropical forest)
 - analysis of how the information is managed (eg software/hardware)
 - analysis of how information is accessed (eg electronic/on-line).

The strength of this analysis is that it should compliment and pick up gaps not covered in the Country Study, and identify biodiversity information which is currently unavailable for national planning. This will provide pointers for biodiversity plans and programmes, and priority areas for institutional capacity building.

An elementary statistical breakdown of findings (eg percentages, means) and graphical presentation of results (eg pie and bar charts, tables) should be used to summarise the information in a preliminary survey report by the Project Team.

Guideline 11

The Project Team should analyse the data obtained, using elementary statistics where necessary, and prepare a preliminary report of the findings of the institutional survey for submission to participants at the Final Workshop (see Section 2.8.1).

2.8 Finalising the Process

2.8.1 Final Workshop

The process and products of the institutional survey should be reviewed at a final workshop, consisting of the Steering Committee, the Project Team, the initial participants in the National Workshop and other important experts/representatives of key institutions during the course of the survey. A preliminary version of the survey report should be made available to participants as a working document on which to comment; this should be provided prior to the final workshop. The workshop should aim to review the findings of the survey, and provide recommendations for building institutional capacity to fill gaps identified in data coverage/biodiversity information management capability in relation to national goals, policies and action plans, and in accordance with obligations as set out in the *Convention on Biological Diversity*. A "case book" of summaries of key institutions' current capabilities and their capacity building requirements may be included as an appendix. The workshop should also decide which institution should be the custodian of the directory of biodiversity information holdings (and metadatabase if constructed), and recommend a process for updating the data holdings to ensure that the initiative of the survey is not lost (see Section 4.3.4). These recommendations and comments should be incorporated in the final version of the survey report.

In addition, the final report should detail aspects of accessibility, networking between institutions, and data flow revolving around the identified "hub" for updating and housing the survey results. Regional and international centres which are of direct relevance to biodiversity information management in the country need to be included in this regard (see Section 5).

Guideline 12

A Final Workshop should be held to review the survey results and comment on the preliminary survey report. Subsequent actions, particularly those required for building institutional capacity for biodiversity information management, should be recommended. The Project Team should then prepare the final survey report.

Guideline 13

Establish a process, including the identification of funding mechanisms, for periodically updating the information gathered during the institutional survey.

2.8.2 Distribution of Final Report and Associated Products

Copies of the final report of the institutional survey, including the recommendations of the final workshop, should be distributed to all participating and interested institutions, including those from other countries. The catalog of data holdings and metadatabase (if prepared) should also be distributed, and consideration should be given to whether the catalog and metadata should be made available to international directories, such as that produced by the Consortium of International Earth Science Information Network (CIESIN) (Burley, 1994), as well as regional information centres and clearing houses. Sources of funding for wide dissemination of outputs will need to be identified.

Guideline 14

Distribute a copy of the final report and associated products to all participating and interested institutions.

2.9 Follow-up

It is suggested that sufficient funds are set aside to finance a Follow-up Workshop to be held in a subsequent year. This workshop would aim to review the effectiveness of follow-up initiatives and the updating mechanism, and monitor developments in the context of national implementation of the CBD. This workshop would also provide the opportunity to assess data exchange, networking, and information flow between institutions following the initial project. How data has been channelled into the decision-making process could also be reviewed.

Guideline 15

Organise a Follow-up Workshop to assess the progress of institutional capacity building, improvements in national biodiversity information management, and effectiveness of the updating mechanism.

3 DEFINING THE SCOPE

3.1 Institutions to Include

A preliminary task of the Steering Committee is to define the scope of the survey (see Section 2.3), ie the **range of institutions** and **types of biodiversity information** to be included. This should be defined while bearing in mind the following points:

- 1) According to Article 2 of the CBD, *biodiversity* is defined as "*the variability among living organisms from all sources including inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems.*" Information relevant to biodiversity management thus includes biological, climatic, geographical, legal, socio-economic and economic data.
- 2) Given practical limits to conducting institutional surveys, it may be necessary to restrict the study to those institutions having a **significant** role in managing biodiversity information. Such institutions would be those that deal with *core* parameters, (eg plants, animals, habitats, ecosystems) but consideration must also be given to other institutions, for example those managing information on physical features (meteorology, geology, soils), economic data (eg eco-tourism), legal aspects of biodiversity management (environmental laws and international treaties), and those institutions responsible for resource management, policy development and decision-making.

It is suggested that the attached matrix (Figure 3.1) is used as an aid to defining the scope of the survey. Two versions of the matrix are provided, one filled in with general institution categories (eg Local Government, Research/Educational) supplied in the text, and the other left blank (see Annex 3). The latter may be copied and filled out in the following way:

- enter the names of biodiversity-related institutions in the left-hand column
- check off the remaining boxes on each row according to the type of biodiversity information held by the institution.

The Steering Committee may wish to send this matrix, with a covering letter explaining the project, to a range of key contacts to assist in identifying relevant institutions managing biodiversity information. The Steering Committee would then be responsible for prioritising those institutions managing biodiversity information for the full institutional survey.

Guideline 16

Use a "matrix approach" to identify the institutions to be covered in the survey, and prioritise if necessary.

3.2 Typical Institutions to Survey

WCMC and IUCN recently completed a study entitled *Availability of Biodiversity Information for East Africa*. A list of the in-country institutions identified as holding or managing data relevant to biodiversity conservation and sustainable use is included in Annex 6; this may serve as a guide to the kinds of institutions that a national survey could target. However, it is important to note that socio-economic, anthropological, eco-tourism related, and legal institutions are not well-represented on this list as a result of resource constraints.

4 CATALOGUING BIODIVERSITY INFORMATION HOLDINGS

4.1 Approach

A major output of the institutional survey is a systematic catalog of which institutions hold what information, and in what format. As described previously these data are collected by questionnaire or other means to consolidate the information into a standard format, facilitated by site visits and/or national/regional/local workshops (see Conduct of the Survey, Section 2). The questions in Part 2 of the questionnaire (see Annex 2) covering "Dataset Information" are the most relevant to the cataloguing process.

Guideline 17

Data on biodiversity information holdings should be collected using a standard information collection form via direct (eg site visits, workshops), and/or indirect (eg posting of questionnaire) approaches.

4.2 Compilation and Presentation of Data

It is likely that a considerable number of biodiversity datasets will be identified in the institutions surveyed. To facilitate the processing and extraction of information it will be necessary, at a minimum, to produce a hard-copy directory of findings (see Section 2.6).

Guideline 18

Produce a hard-copy catalog or directory of biodiversity information holdings (datasets) identified.

4.2.1 Production of Directory

The project *Availability of Biodiversity Information for East Africa* used the "merge" facility in a word processing software package to incorporate dataset descriptions into a metadatabase. Merge files are able to combine specific information entered as records in one file with general information entered in a separate file (for example, layout information). The merging of the file containing records with that containing the layout information can be used to produce a directory of findings, generating a separate sheet(s) for each institutional biodiversity dataset identified. This largely automates and greatly simplifies directory production. The choice of software package used in each country, however, should depend on the availability of in-country support and the systems most widely used nationally.

Guideline 19

Use a standard software package (such as a word processor) to compile and format survey responses electronically to produce a directory of nationally-available biodiversity datasets.

4.2.2 Directory Contents

The hard-copy version of the institutional survey is essentially a reproduction of survey responses. Each response should be reproduced as a separate entry (ie one per institutional dataset held). The directory should contain the following elements:

- table of contents
- introduction giving an overview of the aims of the institutional survey

- explanation of how to use the directory
- catalog of institutional data holdings, ie details of each institutional dataset
- alphabetical listing of institutions surveyed
- glossary
- index.

A directory entry from *Availability of Biodiversity Information for East Africa* is included in Annex 1 as an illustration of a potential directory template.

Guideline 20

The template in Annex 1 can be used as a guide to structuring individual directory entries. The directory contents outline suggested above can be used as a guide to structuring the directory itself.

4.3 Metadatabase Construction

In order to facilitate the national and international exchange of datasets and assist in querying the metadata obtained during the institutional survey, countries may decide to construct an electronic metadatabase of their biodiversity information holdings. This should be viewed as an adjunct to the institutional capacity survey, not necessarily an integral part of it. Many such initiatives are currently being developed around the world, and countries are strongly advised to follow the recommendations below if they wish their metadata to be nationally and internationally compatible.

Guideline 21

The creation of a permanent updateable electronic catalog of national biodiversity datasets is recommended. If such a metadatabase is constructed, Guidelines 22 through 27 should be followed.

4.3.1 Metadatabase Contents

The model questionnaire provided in Annex 2 of this document is based on the metadata format standard that WCMC is establishing (in collaboration with other international organisations) for the recording and exchange of environmental metadata. The standard defines the metadata of importance relating to the **data** that the institution manages.

The model questionnaire contains questions on the **information management capacity** of the institution, found in the "Institutional Resources" section of Part 1 of the questionnaire. Answers relating to information management capacity are of importance for the institutional survey, but are not usually included in the metadatabase. As a rule, the metadatabase concerns itself with the actual **information** that an institution manages (and very brief background details about the institution such as its address and title) and not with the resources that it uses to manage that information.

Guideline 22

Restrict the contents of a metadatabase to a description of institutional datasets and institutional summary. The inclusion of data on institutional resources is not usually incorporated.

4.3.2 *Metadatabase Formats*

In order to facilitate the exchange of data, it is important that the metadatabase format (ie the structure of data entries) and terminology conform to internationally accepted standards. This will facilitate both cooperation and data-sharing between similarly-oriented national institutions and their counterparts in neighbouring countries, and with organisations with international (as opposed to national) scopes. An outline of WCMC's proposed metadatabase formats is given in Annex 7. Further details of currently accepted metadatabase formats and that proposed by WCMC are given in accompanying *Guidelines for Information Management* (Document 3). If a metadatabase is to be developed, metadatabase formats should be discussed at the outset of the project and considered in the context of how the survey is to be conducted (eg using written questionnaire, informal group meeting etc).

Guideline 23

Care should be taken to ensure that metadatabase formats are compatible with international standards in order to promote data-sharing potential. The proposed metadatabase format outlined in Annex 7 can be used as a guide.

Guideline 24

The decision to develop an electronic metadatabase should be taken early on in project development and considered in tandem with the survey approach(es) to be used.

4.3.3 *Metadatabase Data Entry*

The questionnaire results can be directly inserted into a Relational Database Management System (RDBMS) application that has been designed and built on the basis of the metadata standards defined in Annex 7. Database tables may be constructed for each of the entities *Institution*, *Dataset* and *Member*, with table fields corresponding to the attributes as defined in the format definition. WCMC have built their own metadatabase using the INGRES system, but any RDBMS could be used to build a similar system.

For the exchange of metadata with other organisations (eg other NBUs), a metadata exchange format is provided in Annex 7. This is a definition of metadata contents, format and syntax, enabling the latter to be unambiguously recorded.

The syntax used is based on that defined by the US National Aeronautics and Space Administration (NASA) in its *Directory Interchange Format* (DIF) (World Data Center, 1991). This syntax allows the contents of the metadatabase to be output into a standard electronic text report which can be unambiguously understood by other organisations adopting the same format. By following the syntax precisely (ie putting in field titles and values exactly as defined), it becomes possible for an organisation to "auto-ingest" metadata. Thus, rather than the metadata exchange being manually typed into the system, the electronic metadata file can be read in automatically.

Guideline 25

Data should be entered into the metadatabase using definitions provided in Annex 7.

4.3.4 *Metadatabase Updating*

The institution responsible for maintaining the outputs of the institutional survey (see Section 2.8.1), should appoint a metadatabase manager with the explicit duty of keeping metadata current and accurate. The duties of the manager will focus primarily on maintaining the contents of the three core levels up to date (*Institution, Dataset and Member*). An appropriate update period is one year, the procedure commencing at a specified time each year.

Keeping *institutional* metadata up to date is not an onerous task, often requiring little more than simple checks (possibly over the telephone) with relevant dataset managers. New institutions that appear or are "discovered" will need to be dealt with using the original questionnaire/site visit approach, before being added to the country metadatabase.

The *Dataset* and *Product* levels are best handled by printing the contents of the metadatabase, sending these to the individual institutions, and asking them to verify their datasets and products and to supply information on modifications/additions in questionnaire format.

Guideline 26

Ensure that a mechanism is established for regular updating of the metadatabase.

4.3.5 *Metadatabase Documentation*

Whichever metadatabase format is decided upon, clear, concise documentation is essential. This should be produced in the form of a "User's Guide" which should give the technical specifications of the metadatabase and instructions on how to enter records and perform simple searches.

Guideline 27

Produce a User's Guide to assist users of the metadatabase.

5 ASSESSING INSTITUTIONAL LINKAGES

5.1 Approach

A significant output of the survey is enhanced understanding of the **flow** of biodiversity information between institutions, within and outside the country. It is clear that many different types of data are needed for decision making concerning the conservation of biological diversity and sustainable use of its components. A single institution will not have the resources, expertise or mandate to collect, manage or assess all the required data. Thus, it is important to identify both where and how the different sources of information are brought together, and what barriers might exist to the integration process.

By careful comparison with the *Data Flow Model* (Document 1), the linkages between institutions may be compared to the conceptual framework required in support of the CBD. This exercise may reveal collaborative deficiencies which hinder the integrated biodiversity programmes required under the Convention.

5.2 Linkage Information Collection

For any institution, the first level of information required is:

- from which institutions (within and outside the country) are data obtained
- to which institutions are data supplied
- with which institutions are there linkages which do not involve **data** transfer.

The third category above refers to linkages which involve the sharing or exchange of expertise, technology, policies, strategies, and so on, as opposed to the exchange of biodiversity data.

It is possible to obtain this information by means of a questionnaire, such as part 1, Section 3 of the model questionnaire in Annex 2. It is useful to divide linkages into institutions **in** the country and **external** institutions. These linkages, particularly the third group, may be difficult to define initially, and may not all be known to a single individual in the institution. For this reason, a better picture can probably be formed via group discussions during site visits conducted by the Project Team (see Section 2.5.5).

In combination with discussions, the questionnaire approach should result in a list of linked institutions in the form of a table as follows:

Institution Name:

Linked Institution	In-country/ External	Data Provider	Data User	Non-Data Linkage
Institution "W"	IC		M	
Institution "X"	IC	M	M	✓
Institution "Y"	IC	M	m	
Institution "Z"				✓
Institution "1"	E	M		

Legend: IC = In-country
E = External
M = Major
m = Minor

Table 5.1 - Table of Institutional Linkages

The columns could be simple "check marks", but it would be preferable to distinguish between major or frequent data providers (and users) and minor or infrequent providers (and users). In the example of Table 5.1, the notation of **M** for major and **m** for minor has been used.

For each linkage it is also desirable to record the following details:

- administrative arrangements for movement of data (eg informal or formal, are charges made)
- an indication of volume and frequency of data transfer
- physical format of the data transferred (eg documents, maps, computer media, electronic)
- if by computer media what format or standard is used.

For this type of detail there is likely to be such a wide variation that a questionnaire design to elicit meaningful information is difficult. These details are best obtained through discussions with the responsible staff member during a site visit and noted in narrative form. The above points give a general indication of the topics to be covered.

Guideline 28

Aided by a questionnaire or check-list, tabulate in a standard format the data and non-data linkages between the surveyed institution and other national and external institutions.

5.3 Information Compilation and Presentation

5.3.1 Institutional Level

The first level of information supplied will enable the compilation of lists or tables such as Table 5.1, indicating whether the institution is a provider, a user, or both. Note that if any national institution is listed it should have been included in the survey, ie the responses received may indicate omissions in the original scope of institutions surveyed and any such omission should be rectified.

A diagrammatic way of presenting linkage information concerning an institution is also useful. In Figure 5.1, institutions are represented by boxes containing the institution name. Inflows of data from data providers are depicted as arrows pointing towards the surveyed institution, whilst arrows directed towards external institutions represent outflows of data to user institutions. Lines without directional arrows represent non-data linkages. Line types can be used to indicate the degree of linkage. For example, dashed and solid lines can be used to represent minor and major rates of data transfer respectively.

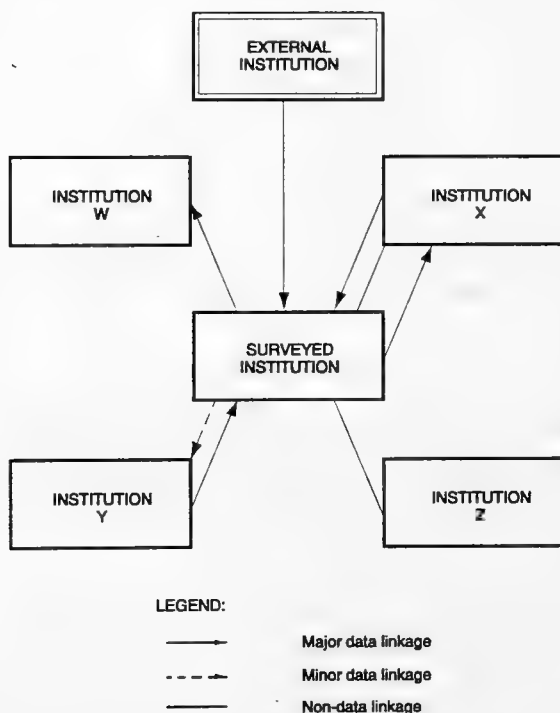


Figure 5.1: Institutional Linkages Diagram

In the example of Figure 5.1, institutions W and X are major users of data; Y is a minor user of the surveyed institution's data; institutions X and Y are major providers of data; and the surveyed institution has non-data (eg policy or technology) linkages with institutions X and Z. In addition, data is received from an external institution (a separate symbol may be

used to distinguish external institutions).

Guideline 29

Identify the linkages of each institution in diagrammatic form using the model of Figure 5.1.

5.3.2 National Level

The diagrammatic approach gives a clear picture of each institution’s linkages from the perspective of the surveyed institution. There will be one such diagram per institution. To usefully analyze the information into a **national** perspective it is necessary to consolidate the information into one overall view. One approach to this is to summarise linkages in a cross tabulation, such as appears in Figure 5.2. Institutions are listed on both axes, with column entries illustrating which institutions are identified as Suppliers of data, and row entries illustrating those Receiving data.

Figure 5.2 shows that institution A supplies data to institutions B, C and D (3 S-entries in the first column), and that it receives data from the same three places (3 R-entries in the first row); institution B supplies data to no-one but receives data from institution A; and so on. Note that each cell in the matrix should be either blank or have two corresponding "Receives/Supplies" entries (the institutions involved in the data transfer having both identified the interaction). This permits cross-checking of the information obtained, and possible follow-up questions to mismatched institutions.

In the sample of Figure 5.2 there are two such anomalous situations: institution C indicates that it is a user of D’s data, but institution D does **not** indicate that it provides data to C, and institute A apparently receives data from institute B, but the latter does not identify the provision of data to A.

Receivers	Suppliers					
	Inst A	Inst B	Inst C	Inst D
Inst A		R	RS	SR		
Inst B	SR					
Inst C	RS			R		
Inst D	SR		SR			
...						
...						

Legend: S: institute in this column is a **supplier** (provider) of data to the institute in this row
 R: institute in this row is a **receiver** (user) of data from the institute in this column

Figure 5.2 National Institute Linkage Cross-Tabulation

The resulting tabulation can help identify which institutions are primary suppliers (many entries in the column for that institution) and which might be primarily data integrators (many entries in the row for that institution).

Again a graphic representation can be very expressive. A convention similar to that in Fig 5.1 above may be used, with the addition of a dot indicating which institution identified the linkage, and arrows depicting the direction of data flow. Each linkage (line) should be identified by both institutions (a dot at both ends of the line). The example shown in Figure 5.3 is based on the same scenario as was the cross-tabulation in Figure 5.2, with the addition of institution E which has non-data linkages to A and a minor data exchange with B. Note the two cases of unmatched dots. Such diagrams will provide a good overall picture of biodiversity information management in the country, and be an aid to identifying missing and inappropriate linkages.

Guideline 30

Summarise national biodiversity information linkages in tabular and diagrammatic form as indicated in Figures 5.2 and 5.3. Clarify and rectify anomalous situations identified in the diagrams.

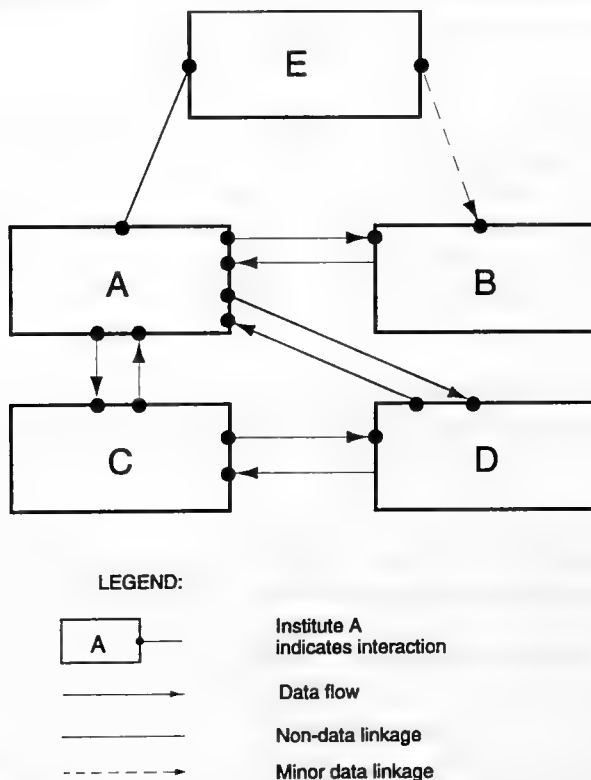


Figure 5.3: National Institute Network Diagram

In the course of discussions with an institution, it may become apparent that some linkages are desirable but have not been achieved. There are a number of possible reasons for these missing links including:

- access procedures and regulations
- security and legal concerns
- charges for data
- lack of knowledge of how to access data
- lack of technology to access data.

These desirable linkages should be identified and tabulated along with a discussion of the barriers to their formation.

Guideline 31

Identify previously unknown linkages, unachieved but desirable linkages, plus the nature of the barriers to linkage formation, as part of the site visit process.

Note that the emphasis in this section is strictly on **linkages** and little is revealed as to the **content** of the data transferred. This, by implication, is best determined through the process of cataloguing information holdings described in Section 4.

6 ASSESSING INFORMATION MANAGEMENT CAPABILITIES

6.1 Approach

To assess the information management capabilities of an institution, basic information on the technical, human and financial resources are needed. Further, there is a need for a general picture of how these resources come together to deliver "information management capability". Two institutions may have similar funding, equipment and numbers of staff, but their capabilities may be very different because of mandate, workload, degree of specialisation, or administrative practices. Some quantitative analysis may be possible, but subjective analysis is also suggested, using information on institutional resources, how data is used in the institution, and the portfolio of projects accomplished.

6.2 Data Collection

The information required falls into two main categories, related to the level of technology in use and the expertise level of staff.

6.2.1 Human Resources

An institution will fail to perform its role without well qualified staff, and thus the most important determinant of its overall capacity is the quality of its employees. For a more thorough discussion of human resource issues in capacity building, refer to the accompanying Document 3.

An assessment of human resources can be made during the survey by quantifying the number of staff, particularly professional staff, and their level of training and expertise in their respective fields. Special attention should be paid to qualifications and skills in applied information technology. Both academic training (degrees and further degrees) and on-the-job experience should be included. Questions of this nature are also incorporated in the Institutional Resources section of the model questionnaire (see Annex 2).

Guideline 33

Obtain basic information on the number and quality of the human resources of the institution using a simple questionnaire design augmented by discussion during a site visit or meeting.

6.2.1 Technical Resources

A preliminary sketch of the institution's technical resources can be accomplished as part of the survey process. Institutions should be asked to provide information on:

- number and type of computer hardware (eg personal computers, work stations)
- operating systems in use (eg DOS, Windows)
- software relevant to biodiversity information management (eg GIS, desk-top mapping, modelling, statistical analysis, spreadsheet, word processing, and database management packages)

- electronic communications (from local networks to "global" systems such as electronic mail and other Internet services).

Suggested questions to elicit this information are included in the Institutional Resources section of the model questionnaire (see Annex 2).

Guideline 32

Obtain basic information on the hardware, software and communications assets of the institution using a simple questionnaire design augmented by discussion during site visits and/or workshops.

Note that it is not necessary to compile a detailed inventory specifying each piece of equipment. General totals by classes are sufficient, such as:

- 3 IBM-PC 386
- 5 IBM-PC 486
- 2 Sun workstations
- 1 A3 flatbed scanner
- 1 A3 inkjet plotter

Again, software can be summarised by listing the names of commonly used packages without attempting an exhaustive list. Suggested headings are:

- word-processors
- spreadsheets
- statistical and modelling packages
- database management systems
- graphics packages
- GIS and desk-top mapping packages
- other applications software.

It may often be the case that detailed software and hardware inventories are already available as routine administrative products in the institution. If that is the case these will suffice, subject to confirmation during the site visit or meeting.

6.2.3 Financial Resources

The financial resources of an organisation, particularly those resources which can be applied to biodiversity information management, are an important factor in determining capability. This is an area which should have been addressed in some detail in locations where a Country Study has been completed, since complete guidelines for evaluating investment in biodiversity are included in the Country Study Guidelines document. However, it may still be necessary to estimate the financial resources available for biodiversity information management in the institution, and this may be quite difficult, depending on the particular accounting structure employed. It is therefore best to approach this problem via discussion during a site visit, and perform an approximate assessment based on these discussions.

Guideline 34

Assess level of institutional financial resources available for biodiversity information management via discussion with senior officials of the institution.

6.3 Information Compilation and Presentation

6.3.1 Institutional Overview

In the model questionnaire (see Annex 2), data on the computer technology used by an institution and the level of its human resources are compiled into a series of summary tables and checklists. To effectively summarise the information for an institution (ie cover the range of questionnaire results), a series of similar tables may be necessary, each being refined following site visit discussions. Some small improvements can be made to assist interpretation, for example:

- software checklists may be modified in light of the responses, ie common packages listed under "other" might be shifted into more specific categories
- the number of staff at a given level of expertise could be entered as a percentage in a further column.

As indicated in the introduction to this section, non-quantitative (subjective) assessments are important. For this reason, a summary report should also be prepared, based on analysis during the site visit. This report would consider the special strengths of the organisation in relation to national biodiversity information strategy, particular attention being made to what specialised information management skills are possessed which might be of value to other partners; the existence of well documented exemplary projects; and possible contributions in terms of the sharing of data and expertise (see Annex 3 for a model table of contents).

Guideline 35

Prepare a report for each institution outlining organisational strengths, exemplary projects, and possible contributions to the national biodiversity information system.

The process of undertaking an institutional survey, especially where site visits feature prominently, provides a good opportunity to identify areas of weakness in institutional capacity, among these the need for additional human resource development and technology. While it is not the primary purpose of the survey, these observations on the needs for capacity building should be recorded during the process, and will provide a useful basis for developing institutional strengthening strategies (see also Section 2.5.2).

Suggestion

Report in narrative form major institutional requirements for capacity building, particularly in terms of human resource development and technology transfer.

6.3.2 National Overview

A summary of the above information is required for national overview. One possibility is to establish a simple database into which the institutional descriptions are entered. This information could then be presented in a variety of different ways as follows:

- to determine the total information management resources of all institutions surveyed
- to tabulate which institutions use which hardware and software
- to tabulate the type of resources in the different institutions.

Such a database (refer to Section 4.3.1) could reveal common practices in software usage (de facto standards), institutional strengths and gaps, and also facilitate answers to *ad hoc* queries. Further, background information such as institutional function and project portfolio could be integrated with the resource database for more general assessments (see 6.4 below).

Suggestion

Maintain basic institutional resource information in a simple, easily updated, database.

6.4 Assessment

The existence of adequate hardware, software, and human resources does not necessarily guarantee capability. An assessment of the latter therefore requires an analysis of institutional competence, by inspection of projects accomplished, reports and papers published, and so on. This can be achieved best through the completion of a site visit and review of the institutional summary report.

The institutional resource database described above may be used to group institutions with respect to their potential for information management capability. Three broad levels (high, medium and low) are suggested for each human or technological resource area. In terms of human resources, "high" may imply a sufficient number of highly qualified staff broadly covering the discipline, and "low", an insufficient number of staff with specific skills only. In the technology area, "high" may imply many networked computers and peripheral devices, and "low", non-networked, obsolete, or absent technology. The meaning of these levels will vary considerably from country to country and are intended purely to assist with internal institutional assessments. The relative role played by the institution in terms of information transfer (as determined by the analysis of linkages (Section 5) is also a key factor.

Institutions with high ratings in both human and technological resources, will probably be key centres in a national biodiversity network, and should thus be given priority in follow-up support. Nevertheless, attention should also be paid to institutional objectives, evidenced by the nature of projects undertaken. These factors may reveal valuable capabilities in organisations which appear to have low-to-medium resource levels.

As capacity building efforts increase, institutional strengths may change rapidly. For this reason, it is important to update institutional information regularly after the completion of the initial survey. Before the survey team is disbanded, a process should therefore be put in place to achieve this, for instance by means of annual inputs of data from participating institutions. Scheduling and funding mechanisms will require careful attention. Annex 5 provides a model table of contents for the final national report.

Guideline 36

Summarise overall national institutional strength by reviewing both quantitative and qualitative information gathered in the survey. Use this information to identify the key national institutions for biodiversity information management.

Guideline 37

Implement a process to keep institutional information up-to-date.

7 REFERENCES

Burley, C. 1994. *CIESIN Metadata Entry Form Instructions*. Consortium for Earth Science Information Network.

Pinburg, U. 1992. *Catalogue of Data Sources (CDS) for the Environment: Analysis and Suggestions for a Meta-data System and Service*. European Environment Agency (EEA).

UNEP 1990. *INFOTERRA Thesaurus of Environmental Terms*, 3rd edn. United Nations Environment Programme, Nairobi.

UNEP 1993. *Guidelines for Country Studies*. United Nations Environment Programme, Nairobi.

WCMC 1993. *Availability of Biodiversity Information in East Africa*. World Conservation Monitoring Centre.

ANNEX 1: SAMPLE DIRECTORY ENTRY

Makerere University Faculty of Science, Zoology Museum

Zoology Department
Makerere University
PO Box 7062
Kampala
Uganda

Tel No:
Fax No: +256 41 530412 (c/o
Wandegeya Post Office)
E-mail:
Telex:

Function of unit/institution: Department: has operated as a depository for specimens used in teaching and for practical demonstration. It is now being developed as a reference point and national repository for animal specimens.

Title or subject of dataset *Mammals*

1 DESCRIPTION OF INFORMATION HOLDINGS

Title or subject of dataset:	Mammals.
Information manager/contact name:	Mr Robert M. Kityo, Museum Curator.
Form of dataset:	Physical specimens/other.
Size and description of holdings:	1,543 physical specimens; a printout of some Ugandan Mammal specimens (1,197 records) in holding at Field Museum of Natural History (Chicago).
Objectives of dataset:	Has functioned as a teaching collection. But is now being developed as a reference source for the National Biodiversity Data Bank of Makerere University.
Localities covered:	Uganda.
Countries covered:	Uganda.
Biomes covered:	Tropical humid forest/tropical dry woodlands/ mountain and highland.
Ecosystems covered:	Savanna/rainforest.
Description of information held:	The bulk of the information is in the form of dried skins and/or skeletal materials. These however also mostly cover small sized animals. Mammals are also stored as pickled specimens.
Completeness, limitations and gaps:	Information not exhaustive, it reflects research interests of people that have gone through the department. The mammals collection has a good representation of the rodent species in the country but not for other groups. The materials are mostly taken from areas that were easily accessible by the researchers.

2 INFORMATION MANAGEMENT

Where dataset is located:	Zoology Department Museum, Makerere University, Kampala, Uganda.
How dataset was acquired/built:	1) Local collection of specimens and mammal materials by the academic staff and students. 2) Through active efforts of collecting specimens by the manager from different parts of the country.
How information is managed:	Card file/catalogued/other. Over 90% catalogued; in the process of computerising the whole collection.
If computerised - hardware:	
 operating system:	
 software:	
Date information collected:	1961-1970, 1971-? 1989-1993.
Are data being actively maintained:	Yes.
Are data part of an ongoing project:	Yes.
Details of project/contract:	National Biodiversity Data Bank, Makerere University Institute of Environment and Natural Resources.

3 ACCESS

Access conditions:	Limited.
Further details on access conditions:	Limited access because it is not a widely known dataset outside of the Zoology Department. Collection is open to students of the University and interested outsiders; access by appointment. Museum Curator.
Outside access through:	
Further details on outside access:	
Documentation of information holdings:	At present only catalog, card index and specimen cards. Plans for future on-screen documentation within 1.5 years.

ANNEX 2: MODEL QUESTIONNAIRE: NATIONAL INSTITUTIONAL SURVEY

DOCUMENT 1: INSTITUTIONAL INFORMATION

1 INSTITUTIONAL DETAILS

Name of institution and Acronym:

Name of department or unit and Acronym:

Type of institution (eg Government, NGO):

Postal address:

Physical address (if different from above):

Telephone no:

Fax no:

E-mail:

Telex:

Core business of unit/institution:
(eg capacity building, research)

Main current activities/projects:

Keywords:

Contact person and status:

Date questionnaire completed:

2 INSTITUTIONAL RESOURCES

Human resources:

Staff total - specify number of core/total

Breakdown of professional staff by level of education/experience and by discipline
(See Table 1)

Areas of outstanding institutional expertise:

Areas of outstanding individual expertise (give names):

Human resource needs:

Computer hardware:

MACHINE TYPE	TOTAL
IBM-PC (personal computer) Specify level (eg 386, 486)	
Apple Macintosh	
Workstation Specify manufacturer and model (eg SUN Sparc IPX)	
Peripherals (eg digitising tables/plotters, printers, other) Specify type	

Computer operating environment:

- DOS
- Windows
- Macintosh
- UNIX
- Other - specify

Computer operating environment needs:

Applications software:

Database Management System (DBMS) - list those in use (eg FoxPro, dBase)

Geographic Information System (GIS) - list those in use (eg ARC/INFO, IDRISI, MapInfo)

Word processing - list those in use (eg WordPerfect, Word)

Statistical and modelling packages - list those in use (eg SPSS, SAS)

Graphics packages - list those in use (eg Harvard Graphics, CorelDraw)

Other applications - list those in use

Application software needs:

Electronic communications:

- Local Area Network (LAN)
- Wide Area Network (WAN)
- E-mail via Internet or commercial provider (eg CompuServe, FidoNet)
- Full Internet connection (eg World Wide Web, Telnet, Ftp, WAIS, Usenet)

Electronic communications needs:

Any other user needs:

3 INSTITUTIONAL LINKS

Up to 10 institutions from which data are obtained regularly:

NAME	NATIONAL	EXTERNAL

Up to 10 institutions to which data are regularly provided:

NAME	NATIONAL	EXTERNAL

*Up to 10 institutions with which non-data linkages are established:
(eg exchange of expertise, policy or advice)*

NAME	NATIONAL	EXTERNAL

Up to 10 potential collaborating organisations:

NAME	NATIONAL	EXTERNAL

DOCUMENT 2: DATASET INFORMATION

1 DESCRIPTION OF DATASET (complete one copy for *each* dataset)

Note: For purposes of this survey, a "dataset" can be a single file, but more commonly will consist of a collection of thematically related files which can be conveniently described as a group, such as "Forest ecology data from the XX region". It is **not** intended that each individual record, map, report, GIS coverage or small electronic file be separately listed and described.

1.1 Name or subject area of dataset:

1.2 Objectives of dataset:

1.3 Keywords:

1.4 Name of information manager:

1.5 Form of dataset:

Size of information holdings:

- Bibliographic collections
- Physical specimens (eg herbarium sheets)
- Field records
- Maps
- Tables
- Reports
- GIS holdings (eg national wetlands coverage)

- Database information (electronic records)
- Word processing files
- Remote sensing data (specify sensor type, eg AVHRR
MSS, TM, SPOT, aerial photography)
- Other - specify

1.6 Location of dataset:

1.7 How the information is being managed:

- All computerised
- Partly computerised (%)
- Card file
- Catalogue
- Other - specify

1.8 How the dataset was acquired/built:

1.9 If computerised, software used to manage dataset:
(eg Excel, dBase, FoxPro, Ingres, Oracle, ARC/INFO)

2 INFORMATION COVERAGE

2.1 Area of covered (provide a map where possible):
(eg local/regional/national/international/global)

2.2 Themes covered:

2.3 Maintenance of the information:

Active (yes/no):

Part of an ongoing project/contract (yes/no/details):

2.5 Time period that information covers:
(eg 1914-1949)

Start day/month/year:

Stop day/month/year:

2.6 Completeness, limitations, and gaps in the information:

3 INFORMATION ACCESS

3.1 Conditions:

No outside access/use

- Limited access (please give details)
- Freely available (at any time?)
- On payment of funds (eg cost of recovery; commercial sale)
- Other - specify

3.2 Outside access through:

- Published material
- Diskette/tape
- On-site
- On-line - specify network(s)
- Other - specify

3.4 Documentation of information holdings:
(eg descriptions available, user guides, on-screen documentation)

3.5 Other organisations known to hold biodiversity information:
(names and addresses)

Title and position of person completing this form:

Please return as **quickly as possible** to:

ANNEX 4: MODEL TABLE OF CONTENTS: INSTITUTIONAL REPORT

1 Institutional Details

2 Biodiversity Dataset Holdings

2.1 Description of Dataset(s)

2.2 Information Coverage

2.3 Information Access

3 Information Management Capabilities

3.1 Technical Resources

3.2 Human Resources

3.3 Financial Resources Available for Biodiversity Information Management

3.4 Organisational Strengths and Competencies

4 Institutional Linkages

4.1 Data Supply and Provision, and Non-data Linkages

4.2 Gaps and Weaknesses

5 Needs Assessment (optional)

6 Institutional Summary

ANNEX 5: MODEL TABLE OF CONTENTS: FINAL NATIONAL REPORT

PART 1: SUMMARY OF FINDINGS

- 1 Introduction**
 - 1.1 The Project Context
 - 1.2 Chronology of Project Development and Implementation
 - 1.3 Methodology Adopted

- 2 Analysis and Summary of Results**
 - 2.1 Biodiversity Dataset Holdings: Coverage, Gaps, Accessibility
 - 2.2 Information Management Capabilities:
 Overview of Technical, Human and Financial Resources Available
 - 2.3 Institutional Linkages: The National Picture
 - 2.4 Needs Analysis (optional)
 - 2.5 State of National Biodiversity Information Management: A Summary

- 3 Project Conclusions**

- 4 Recommendations for Future Action**

- ANNEX 1 Information Brochure**
- ANNEX 2 Questionnaire(s)**
- ANNEX 3 Synopsis of Workshop Proceedings**
- ANNEX 4 Case Book of Selected Institutional Reports**

PART 2: CATALOG OF SURVEY ENTRIES

- 1 Introduction**
- 2 Use of Catalog**
- 3 Explanation of Catalog Headings**
- 4 Questionnaire/Survey Entries**
- 5 Alphabetical Listing of Institutions**
- 6 Glossary**
- 7 Index**

ANNEX 6: TARGET INSTITUTIONS IDENTIFIED IN EAST AFRICA

KENYA

African Centre for Technology Studies
African Wildlife Foundation
Department of Resource Surveys and Remote Sensing
Douglas-Hamilton Associates - Ecological Consultancies
East African Natural History Society
East African Wildlife Society
Game Ranching Ltd.
Genebank (KARI)
International Centre for Research in Agroforestry (ICRAF)

KENGO

Kenya Forestry Research Institute (KEFRI) - Entomology
Kenya Forestry Research Institute (KEFRI) - Pathology and Mycology
Kenya Marine and Fisheries Research Institute (KMFRI) - Computer Section
Kenya Meteorological Department - Data Processing Section
Kenya Trypanosomiasis Research Institute
Kenya Wildlife Service (KWS) - GIS Unit
Kenya Wildlife Service (KWS) - Scientific Services, Computer Department
Kenya Wildlife Service (KWS) - Veterinary Department
National Council for Science and Technology - Biological and Physical Sciences
National Environment Secretariat
National Museums of Kenya - Centre for Biodiversity
National Museums of Kenya - Coastal Forest Conservation Unit
National Museums of Kenya - Department of Herpetology
National Museums of Kenya - Herbarium
National Museums of Kenya - Invertebrate Zoology
National Museums of Kenya - Indigenous Food Plants Programme
National Museums of Kenya - Kenya Indigenous Forest Conservation (KIFCON)
National Museums of Kenya - Mammalogy Department
The Kora Research Project
Wildlife Conservation International
World Wide Fund For Nature (WWF) East African Regional Office

TANZANIA

Animal Disease Research Institute
College of African Wildlife Management
Fisheries Division - Ministry of Tourism, Natural Resources and the Environment
Frontier Tanzania - Coastal Forest Research Programme
Horticulture Research and Training Institute - Ministry of Agriculture
Institute of Resources Assessment - University of Dar es Salaam
Institute of Traditional Medicine
Malihai Clubs of Tanzania - Research and Training
National Environment Management Council - Documentation Centre
National Land Use Planning Commission
National Museums of Tanzania

National Tree Seed Centre
Catchment Forestry Project - Ministry of Tourism, Natural Resources and the Environment
Serengeti Wildlife Research Centre
Serengeti Wildlife Research Institute
Sokoine University of Agriculture
Tanzania Commission for Science and Technology - Information Unit
Tanzania Fisheries Research Institute
Tanzania Forestry Research Institute
Tropical Pesticides Research Institute - Agricultural Entomology Section
Tropical Pesticides Research Institute - National Herbarium
Tropical Pesticides Research Institute - National Plant Genetic Resources Centre
University of Dar es Salaam - Library
University of Dar es Salaam - Demographic Training Unit
University of Dar es Salaam - Department of Botany
University of Dar es Salaam - Department of Zoology and Marine Biology
Wildlife Department - Ministry of Tourism, Natural Resources and the Environment

UGANDA

Budongo Forest Project
Department of Environmental Protection - National Environment Information Centre
Department of Environmental Protection -
 National Wetlands Conservation and Management Programme
Forest Department - Natural Forest Conservation Project
Game Department
Forest Department - Nyabyeya Forest College
Institute of Tropical Forest Conservation
Kawanda Agricultural Research Institute
Makerere University - Zoology Department/Museum
Makerere University - Veterinary Microbiology
Makerere University Biological Field Station
Makerere University - Herbarium
Makerere University Institute of Environmental and Natural Resources -
 National Biodiversity Data Bank
Ministry of Agriculture, Animal Resources and Fisheries - Botanical Garden
Ministry of Lands, Housing and Urban Development - Department of Lands and Surveys
Ministry of Public Health - Institute of Public Health
Mount Elgon Conservation and Development Project
Namulonge Agricultural Research Institute
National Agricultural Research Organisation - Fisheries Research Institute
Natural Chemotherapeutics Research Laboratory
Uganda Institute of Ecology
Uganda National Parks/Game Department - Wildlife College, Lake Katwe

ANNEX 7: MODEL METADATABASE FORMAT

For the purposes of conducting an Institutional Survey, WCMC has evaluated a number of different existing environmental metadata formats, primarily the European Environment Agency (EEA) Catalogue of Data Sources (CDS) (Pinborg, 1992), the UNEP Global Resource Information Database (GRID) Metadatabase (UNEP, 1992), the Consortium for International Earth Science Information Network (CIESIN) Catalogue Service (Burley, 1994), and the UNEP Harmonization of Environmental Measurement (HEM) HemDisk (UNEP, 1994). Each of these addresses certain requirements, but also leaves other requirements unfulfilled. The metadata exchange standard proposed here is based on these existing or developing metadatabase systems, supplemented with additions to meet the requirements of biodiversity information management.

A three-layer metadata model is proposed (see Figure A6.1), based on the hierarchy of entities: **Institution > Dataset > Member**. Additional entities can be added to this core as necessary. For instance, some institutions may wish to add a "Personnel" entity in order to record information on individuals in the institution, which then can be linked to the three core entities.

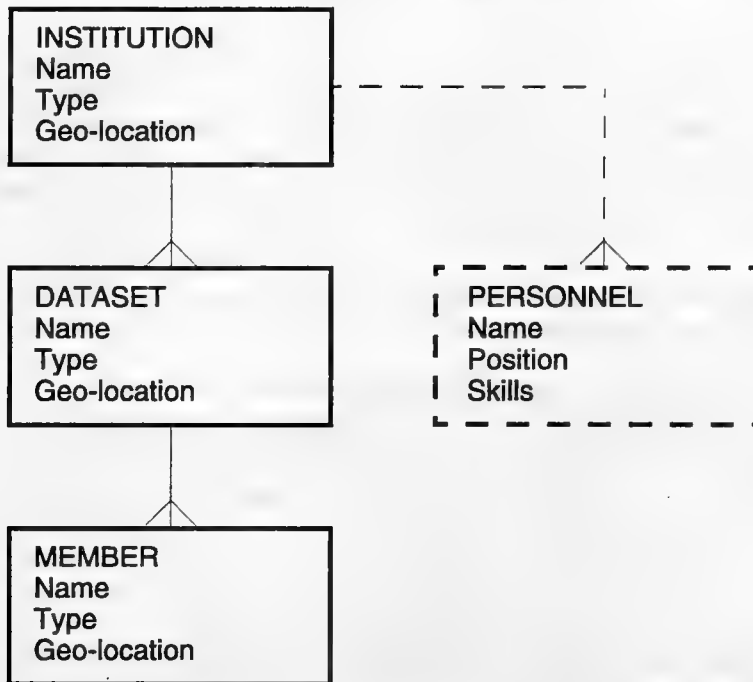


Figure A6.1: Metadata Entity Hierarchy

The **Institution** is defined as a recognisable organisation that stores and maintains information of relevance to the *Convention on Biological Diversity*.

The **Dataset** is defined as a collection of data and accompanying documentation maintained at an **Institution**. A collection of data refers to one or a series of Data Members which relate to a specific *theme* or *geographic region*. A **Dataset** entity occurs under that name in the UNEP-GRID metadatabase (UNEP, 1992), and under the name "Directory Entry" in the CIESIN metadatabase (Burley, 1994).

The **Member** can be regarded as equivalent to a data file, paper report, or map, and must be a component of a particular **Dataset**. Thus **Members** are the lowest level "concrete" entities that can be distributed by an institution.

There are many different types of **Member**, for instance reports, paper maps, GIS coverages, and database files. One way to accommodate such variety is to abandon a structured format and describe members using free text. However, this means that **Member** metadata can have widely varying degrees of completeness and little standardisation. These guidelines recommend the definition of a **Member** format that incorporates as much structure as possible, whilst still permitting all types of **Member** to be recorded.

For the purposes of the institutional survey, information will need to be collected on the top two levels of the hierarchy (**Institution** and **Dataset**). The collection of metadata for the lowest level (**Member**) is not essential for the purposes of the survey and is not included in the questionnaire. This level is provided in the metadata format so that the National Biodiversity Unit (NBU) has the later option of recording and storing metadata at this level of detail (possibly within a metadatabase system).

The remainder of this annex provides sample metadatabase entries and field definitions for **Institution**, **Dataset**, and **Member** metadata.

Example of a completed **Institution** metadatabase entry

Institution_ID:	0001
Name:	World Conservation Monitoring Centre
Acronym:	WCMC
Type:	Non-governmental
Theme:	Information Services
Keywords:	biodiversity; conservation; information
Postal_Address:	219 Huntingdon Road
Postal_Code:	CB3 0DL
City:	Cambridge
Country:	United Kingdom
Contact_Person:	Jo Taylor
Contact_Status:	Information Officer
Telephone:	44-1223-277314
Fax:	44-1223-277136
Email:	Internet: info@wcmc.org.uk

Update_Date: 1994-09-01
Mission: To provide research, information and technical services so that decisions affecting the conservation and sustainable use of biological resources may be based on the best available scientific information.

Field definitions of Institution metadata

Institution_ID:

Definition Internal identification by which the institution is catalogued within the metadatabase. Institution ID numbers are assigned automatically by the system every time a new institution is added. This ID is not used in metadata exchange among organisations and as such is not an official part of the metadata format definition, but part of the metadatabase system definition.

Format Integer of length = 4.
Status System generated.
Example 0001

Name:

Definition Official name of the institution.
Format Maximum 50 characters.
Status Mandatory.
Example World Conservation Monitoring Centre.

Acronym:

Definition Short name of the institution.
Format Maximum 15 characters.
Status Optional.
Example WCMC.

Type:

Definition The organisation type, selected from one of the following: governmental; non-governmental; commercial; academic; inter-governmental; United Nations.
Format Maximum 20 characters.
Status Mandatory.
Example Non-governmental.

Theme:

Definition The primary function of the organisation, selected from one of the following: Research; Consultancy; Information Services; Campaigning. The selection of a primary function keyword is not intended to wholly define the scope of the organisation. Detailed description of the function of the organisation can be expanded on in the "Mission" section.
Format Maximum 30 characters.
Status Mandatory.
Example Information Services.

Keywords:

- Definition* Keyword(s) by which the institution can define its activities and areas of interest. Used for searching by the metadatabase.
- Format* Use terminology lists where possible (eg Pinborg, 1992; UNEP, 1990; WDC, 1991). Several terms may be entered if separated by semicolons.
- Status* Optional.
- Example* biodiversity; conservation; information.

Postal Address:

- Definition* Postal address of the organisation (within city location).
- Format* Maximum 50 characters.
- Status* Mandatory.
- Example* 219 Huntingdon Road.

Postal Code:

- Definition* Official postal code of the organisation.
- Format* Maximum 15 characters.
- Status* Optional.
- Example* CB3 0DL.

City:

- Definition* City location of the organisation.
- Format* Maximum 30 characters.
- Status* Mandatory.
- Example* Cambridge.

Country:

- Definition* Country location of the organisation.
- Format* Use standard United Nations country names.
Maximum 30 characters.
- Status* Mandatory.
- Example* United Kingdom.

Contact Person:

- Definition* Primary contact person or department for further information on the organisation.
- Format* First, Middle and Last Name.
Maximum 40 characters.
- Status* Mandatory.
- Example* Jo Taylor.

Contact Status:

- Definition* Position or status of the contact person within the organisation.
- Format* Maximum 30 characters.
- Status* Optional.
- Example* Information Officer.

Telephone:

Definition Telephone number to reach the contact person.
Format Hyphen-separated codes with all the required numbers for international dialling, ie CountryCode-AreaCode-TelephoneNumber.
 Maximum 30 characters.
Status Optional.
Example 44-1223-277314

Fax:

Definition Fax number to reach the contact person.
Format Hyphen-separated codes with all the required numbers for international dialling, ie CountryCode-AreaCode-FaxNumber.
 Maximum 30 characters.
Status Optional.
Example 44-1223-277136

Email:

Definition Electronic mail address to reach the contact person.
Format Network > email address (note that the network name is required, such as: SPAN, Telemail, Internet, NSI/DECnet, BITNET, OMNET, CompuServe).
 Maximum 50 characters.
Status Optional.
Example Internet > info@wcmc.org.uk

Update_Date:

Definition Date of when this metadata entity was last updated.
Format YYYY-MM-DD.
Status Mandatory.
Example 1994-08-30

Mission:

Definition A brief summary of institution's mission, goals and activities.
Format Maximum 250 characters.
Status Optional.
Example To provide research, information and technical services so that decisions affecting the conservation and sustainable use of biological resources may be based on the best available scientific information.

Example of a completed Dataset metadatabase entry

Dataset_ID: 0036
Name: African Protected Areas GIS
Acronym: AFRICAPA
Theme: Terrestrial ecosystems
Keywords: protected areas; wildlife
Institution_Name: World Conservation Monitoring Centre
Institution_Contact: Ian Barnes

Start_Date:
Stop_Date:
Update_Date: 1992-02-21
Geo_Coverage: Continental
Continent: Africa
Region:
Country:
Sub_Natl:
Minimum_Longitude: -17.50
Maximum_Latitude: +37.10
Maximum_Longitude: +51.30
Minimum_Latitude: -34.80
Abstract_Filename: /gis/docs/afr_pa_gis.wp

Field definitions of Dataset metadata

Dataset_ID:

Definition Internal identification by which the dataset is catalogued within the metadatabase. Dataset ID numbers are assigned automatically by the system every time a new dataset is added. This ID is not used in metadata exchange among organisations and as such is not an official part of the metadata format definition, but part of the metadatabase system definition.

Format Integer of length = 4.

Status System generated.

Example 0036

Name:

Definition The name given to the dataset or activity being described. The title should be descriptive enough to allow the reader to make a reasonable decision as to whether the data may be of interest.

Format Maximum 50 characters.

Status Mandatory.

Example African Protected Areas GIS.

Acronym:

Definition Short name (acronym) for the dataset, if there is one.

Format Maximum 15 characters.

Status Optional.

Example AFRICAPA.

Theme:

Definition This is the theme or parameter being measured by the dataset. The keyword entered is the most general, and should, if possible, be taken from the standard terminology lists.

Format Use INFOTERRA terminology list (UNEP, 1990).

Maximum 31 characters.

Status Mandatory.

Example Terrestrial ecosystems.

Keywords:

Definition Keyword(s) are additional terms that describe and characterise the dataset. Used for searching by the metadatabase.

Format Use terminology lists where possible (eg Pinborg, 1992; UNEP, 1990; WDC, 1991). Several terms may be entered if separated by semicolons.

Status Optional.

Example protected areas; wildlife.

Institution Name:

Definition Primary institutional holding the stored data or coordinating centre of the activity. A link may be set up from this field to the "Name" field of the **Institution** metadata level.

Format Maximum 50 characters.

Status Mandatory.

Example World Conservation Monitoring Centre.

Institution Contact:

Definition Contact person for technical support, information about the technical content of the data, access to the data, and how the data are stored (this field doubles for the NASA DIF *Technical_contact* field (WDC, 1991).

Format First, Middle and Last Name.
Maximum 40 characters.

Status Mandatory.

Example Ian Barnes.

Start Date:

Definition The beginning date for a temporal dataset (eg satellite data collected over a period of weeks, months or years).

Format YYYY-MM-DD.

Status Optional.

Example 1986-06-31

Stop Date:

Definition The ending date for a temporal dataset (see above). This date can also be used for datasets which are periodically updated, such as climatic data. If left blank, the data collection is assumed to be on-going.

Format YYYY-MM-DD.

Status Optional.

Example 1986-08-31

Update Date:

Definition Date of when the dataset was last updated.

Format YYYY-MM-DD.

Status Mandatory.

Example 1994-08-30

Geo_Coverage:

Definition For geo-referenced datasets, this is the coverage area selected from one of the following: global; continental; regional; national; sub-natl.

Format Maximum 15 characters.

Status Mandatory for geo-referenced datasets.

Example Continental.

Continent:

Definition For geo-referenced datasets, this is the name of the continent in which the dataset occurs.

Format Use INFOTERRA terminology list (UNEP, 1990).
Maximum 30 characters.

Status Mandatory if "Geo_Coverage" field is continental, regional, national, sub-natl.

Example Africa.

Region:

Definition For geo-referenced datasets, this is the geographic region in which the dataset occurs.

Format Use INFOTERRA terminology list (UNEP, 1990).
Maximum 30 characters.

Status Mandatory if "Geo_Coverage" field is regional, national, sub-natl.

Example Tropical Africa.

Country:

Definition For geo-referenced datasets, this is the country in which the dataset occurs.

Format Use United Nations standard country names (English language version).
Maximum 30 characters.

Status Mandatory if "Geo_Coverage" field is national or sub-natl.

Example Tanzania.

Sub_Natl:

Definition For geographically-referenced data, this is the name of the sub-national area which is covered by the dataset.

Format Free-text area names. Maximum 30 characters.

Status Mandatory if "Geo_Coverage" field is "sub-natl".

Example Serengeti

Minimum_Longitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Minimum_Longitude* refers to the western-most longitude point covered, in signed decimal degrees.

Format Floating point length = 8.

Status Mandatory for geo-referenced data.

Example -17.50

Maximum_Latitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Maximum_Latitude* refers to the northern-most latitude point covered, in signed decimal degrees.

Format Floating point length = 8.

Status Mandatory for geo-referenced data.

Example +37.10

Maximum_Longitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Maximum_Longitude* refers to the eastern-most longitude point covered, in signed decimal degrees.

Format Floating point, length = 8.

Status Mandatory for geo-referenced data.

Example +51.30

Minimum_Latitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Minimum_Latitude* refers to the southern-most latitude point covered, in signed decimal degrees.

Format Floating point of length = 8.

Status Mandatory for geo-referenced data.

Example -34.80

Abstract_Filename:

Definition The full path name of the file containing the abstract text for the dataset. This is composed of unstructured text containing information about the dataset that cannot be recorded in the other metadata fields. Significant information should be included, such as the objectives of the data collection or study, scope, methodology, and major findings. Also included is a general discussion of the thematic contents of the dataset.

Further fields to consider:

Quality:

Definition Information about any quality procedures followed in producing the dataset or any other indicators as to the quality of the data. This includes a description of any known gaps or anomalies. Indicators include, but are not limited to: frequency of use; processing and quality assurance procedures; frequency of updates; number of peer reviewed articles; and any prior user-feedback.

Format Free text.

References:

Definition Key bibliographic references pertaining to the metadata entry.

Format Use any bibliographic style, but be consistent. The style required by the Journal of Geophysical Research (JGR) is suggested.
Maximum 2000 characters (25 lines, 80 characters per line) free text.

Example Kolenkiewicz, R. and Martin, C. F. 1982, Seasat altimeter height calibration, *J. Geophys. Res.*, 87 (C5), pp.3189-3198.

Example of a completed Member metadatabase entry

Member_ID: 0012
Dataset_Name: African Protected Areas GIS
Name: Serengeti GIS
Type: Raster GIS
Keywords: GIS; ARC/INFO
Start_Date:
Stop_Date:
Update_Date: 1994-08-30
Geo_Coverage: Sub-natl
Continent: Africa
Region: Tropical Africa
Country: Tanzania
Sub_Natl: Serengeti
Minimum_Longitude: +33.80
Maximum_Latitude: -1.45
Maximum_Longitude: +35.20
Minimum_Latitude: -3.15
Final_Process: ARC/INFO export with no compression
Process_Source: ARC/INFO coverage on VAX disk
Access_and_Distrib: Public
File_Size: 3400
File_Medium: Various media
Comments: Digitised from the 1:100,000 map "Serengeti National Park", 1987

Field definitions of Member metadata

Member_ID:

Definition Internal identification by which the member is catalogued within the metadatabase. Member ID numbers are assigned automatically by the system every time a new data member is added. This ID is not used in metadata exchange among organisations and as such is not an official part of the metadata format definition. However, it is part of the metadatabase system definition.

Format Integer of length = 4.

Status System generated.

Example 0012

Dataset_Name:

Definition Name of the dataset to which this data member is attached. A link may be set up from this field to the "Name" field of the **Dataset** metadata level.

Format Maximum 50 characters.

Status Mandatory.

Example World Conservation Monitoring Centre.

Name:

Definition The name given to the data member being described. The title should be descriptive enough to allow a reader to make a reasonable decision as to whether the data may be of interest. The name is a one-line description of a member, and should tell the user what is unique about this particular member/data file that distinguishes it from all other members of the same dataset. It may state the version in a series of data files, a special process, resolution or scale of a data file.

Format Maximum 50 characters free text

Status Mandatory

Example Serengeti GIS

Type:

Definition The type of the member, selected from one of the following: report; tabular data; image; paper map; vector GIS; raster GIS. A 'raster' member can be a raster map or classified (categorised) satellite image, whereas an 'image' member is either raw satellite data or non-categorised, non-discrete data (eg vegetation indices). A 'vector' member can consist of points, lines (arcs) or areas (polygons).

Format Maximum 20 characters.

Status Mandatory.

Example Raster GIS.

Keywords:

Definition Keyword(s) are additional terms that describe and characterise the data member. Used for searching by the metadatabase.

Format Use terminology lists where possible (eg Pinborg 1992; UNEP, 1990; WDC, 1991). Several terms may be entered if separated by semicolons.

Status Optional.

Example GIS; ARC/INFO.

Start_Date:

Definition The beginning date for a temporal data member (eg satellite-based data collected over a period of weeks, months or years). In most cases of non-temporal data, it will be the same as the "Start_Date" at the Dataset metadata level.

Format YYYY-MM-DD.

Status Optional.

Example 1986-06-31

Stop_Date:

Definition The date on which the final process was executed.

Format YYYY-MM-DD.

Status Optional.

Example 1986-08-31

Update_Date:

Definition The date when this metadata entity was last updated.
Format YYYY-MM-DD.
Status Mandatory.
Example 1994-08-30

Geo_Coverage:

Definition For geo-referenced datasets, this is the coverage area selected from one of the following: global; continental; regional; national; sub-natl.
Format Maximum 15 characters.
Status Mandatory for geo-referenced datasets.
Example Sub-natl.

Continent:

Definition For geo-referenced datasets, this is the name of the continent in which the dataset occurs.
Format Use INFOTERRA terminology list (UNEP, 1990).
Maximum 30 characters.
Status Mandatory if "Geo_Coverage" field is continental, regional, national, sub-natl.
Example Atlantic Ocean.

Region:

Definition For geo-referenced datasets, this is the geographic region in which the dataset occurs.
Format Use INFOTERRA terminology list (UNEP, 1990).
Maximum 30 characters.
Status Mandatory if "Geo_Coverage" field is regional, national, sub-natl.
Example SE Asia.

Country:

Definition For geo-referenced datasets, this is the country in which the dataset occurs.
Format Use United Nations standard country names (English language version).
Maximum 30 characters.
Status Mandatory if "Geo_Coverage" field is national or sub-natl.
Example Brazil.

Sub_Natl:

Definition For geographically-referenced data, this is the name of the sub-national area which is covered by the dataset.
Format Free-text area names. Maximum 30 characters.
Status Mandatory if "Geo_Coverage" field is "sub-natl".
Example Serengeti National Park.

Minimum_Longitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Minimum_Longitude* refers to the western-most longitude point covered, in signed decimal degrees.

Format Floating point length = 8.
Status Mandatory for geo-referenced data.
Example -17.50

Maximum_Latitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Maximum_Latitude* refers to the northern-most latitude point covered, in signed decimal degrees.

Format Floating point length = 8.
Status Mandatory for geo-referenced data.
Example +37.10

Maximum_Longitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Maximum_Longitude* refers to the eastern-most longitude point covered, in signed decimal degrees.

Format Floating point, length = 8.
Status Mandatory for geo-referenced data.
Example +51.30

Minimum_Latitude:

Definition For geo-referenced datasets, the four coverage fields indicate the spatial extent of the dataset. *Minimum_Latitude* refers to the southern-most latitude point covered, in signed decimal degrees.

Format Floating point of length = 8.
Status Mandatory for geo-referenced data.
Example -34.80

Final_Process:

Definition A one line description of the last processing step used to create the data member.

Format Maximum 50 characters.
Status Optional.
Example ARC/INFO export with no compression.

Process_Source:

Definition Description of the data file or other product used as input to the final process. Should enable users to determine the exact origin of the final data member.

Format Maximum 50 characters.
Status Optional.
Examples ARC/INFO coverage on VAX disk.
Printed reports sourced from UNEP.

Access_and_Distrib:

Definition A one word description of the dataset specifying its access and distribution rights, selected from: **public** (files in the public domain which can be freely

distributed); **limited** (files in the public domain but not for all users due to format, size considerations etc); **source** (files requiring source permission); **in-house** (files which are proprietary and cannot be distributed at all outside of the data centre).

Format From selection (maximum 10 characters).
Status Mandatory.
Example Public.

File_Size:

Definition The file size for each data member (electronic data files only) in its final storage format, in kilobytes rounded to the nearest integer.

Format Integer of length = 4.
Status Optional.
Example 340

File_Medium:

Definition Type of medium on which the data file is stored at the institution. Include any special characteristics about the storage medium such as media (eg paper, hard disk, magnetic tape), and capacity. It should be noted that the storage medium may differ from that used for distribution.

Format Maximum 50 characters.
Status Optional.
Example 270 Mb Seagate hard disk, 1.44 Mb 3.5" floppy disk, paper report.

Comments:

Definition This field may be used for a variety of purposes, typically to provide details on the data member not accounted for elsewhere. These might include the member's processing history or origins, the reason why the data member was initially acquired or created, the name and scale of a source map, or the projection used for a GIS product.

Format Maximum 250 characters.
Status Optional.
Example Digitised from the 1:100,000 map "Serengeti National Park", 1987.

Metadatabase Updating

Keeping the **Institution** metadata current should not be unduly laborious and can probably be verified by checking the validity of the contents with a phone call to the contact person at the institution. New institutions that appear or are "discovered" will need to be dealt with using the original questionnaire/site visit approach and then adding them to the country metadatabase.

The **Dataset** and **Member** levels are best handled by sending a printout of the contents of the metadatabase to each institution, asking them to verify their entries and supply modifications and additions using the questionnaire format.

Annex 7 References

Pinburg, U. 1992. *Catalogue of Data Sources (CDS) for the Environment: Analysis and Suggestions for a Meta-data System and Service*. European Environment Agency (EEA).

UNEP 1990. *INFOTERRA Thesaurus of Environmental Terms (3rd ed.)*. UNEP, Nairobi.

UNEP 1992. *The Grid Meta-Database (MDb) Entity-Attribute Definitions*. UNEP-GRID, Geneva.

UNEP 1994. *An Introduction to HEM and the HEMDisk* (ed. Crain, I.K.). Office of Harmonization of Environmental Measurement. UNEP, Munich.

World Data Center 1991. *Directory Interchange Format Manual (version 4.0)*. World Data Center, NASA.

ANNEX 8: LIST OF ACRONYMS & ABBREVIATIONS

CBD	Convention on Biological Diversity
CDS	Catalog of Data Sources (EEA)
CIESIN	Consortium for International Earth Science Information Network
DIF	Directory Interchange Format
EEA	European Environment Agency
GEF	Global Environment Facility
GIS	Geographical Information System
IUCN	World Conservation Union
NASA	National Aeronautics and Space Administration (US)
DIF	Directory Interchange Format (NASA)
NBU	National Biodiversity Unit
NGO	Non-Governmental Organisation
RDBMS	Relational Database Management System
UNEP	United Nations Environment Programme
GRID	Global Resources Information Database (UNEP)
HEM	Harmonization of Environmental Measurement (UNEP)
WCMC	World Conservation Monitoring Centre
WOCE	World Ocean Circulation Experiment

NB See also the index of acronyms and abbreviations in the *Resource Inventory* (Document 4).



**WORLD CONSERVATION
MONITORING CENTRE**

**World Conservation Monitoring Centre
219 Huntingdon Road
Cambridge CB3 0DL
United Kingdom**

**Telephone +44 223 277314
Fax +44 223 277136**



The World Conservation Monitoring Centre is a joint-venture between the three partners who developed the *World Conservation Strategy* and its successor *Caring for the Earth*: IUCN-The World Conservation Union, UNEP-United Nations Environment Programme, and WWF-World Wide Fund for Nature.