

**FACULTY OF ENGINEERING AND BUILT ENVIRONMENT**

**ENVIRONMENTAL INFORMATION SYSTEMS:**

**The Development and Implementation of the Lake Rukwa Basin Integrated Project  
Environmental Information System (LRBIP-EIS) Database, Tanzania**

*Submitted in fulfillment of the requirements of the course*

***SUR524W MASTERS DISSERTATION: GEOMATICS***

*for the degree*

***MASTER OF SCIENCE (MSc Eng) in ENGINEERING***

*by*

**Charles Paradzayi**

**Department of Geomatics**

**The School of Architecture, Planning and Geomatics**

*Supervised by:*

**Professor Heinz Rüther**



**UNIVERSITY OF CAPE TOWN**

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

# DECLARATION BY CANDIDATE

I hereby declare that this thesis is my own work, and where appropriate I have acknowledged the work of others to the best of my knowledge. This thesis has not been submitted for a degree at any other University.

Signed by Candidate

Signature Removed

Cape Town

2 January 2003

# ABSTRACT

The quest for sustenance inevitably forces mankind to exploit natural resources found within their environs. In many cases, the exploitation results in massive environmental degradation that disrupts the ecosystem and causes loss of bio-diversity. There is generally a lack of information systems to monitor and provide quantitative information on the state of the affected environment. Decision-makers usually fail to make informed decisions with regard to conservation strategies. The need to provide decision-makers with quantitative environmental information formed the basis of this thesis.

An integrated environmental information system (EIS) database was developed according to the Software Development Methodology for three of the identified environmental sectors. This involved detailed user needs assessment to identify the information requirements (both spatial and textual) for each sector. The results were used to design separate data models that were later merged to create an integrated data model for the database application.

A fisheries application prototype was developed to implement the proposed database design. The prototype has three major components. The Geographic Information System (GIS) handles the spatial data such as rivers, settlements, roads, and lakes. A relational database management system (RDBMS) was used to store and maintain the non-spatial data such as fisherman's personal details and fish catch data. Customized graphical user interfaces were designed to handle the data visualization and restricted access to the GIS and RDBMS environments.

# ACKNOWLEDGEMENTS

Special mention goes to all my family (especially to *Stashia, Audrey, Kudakwashe and Danaï*), for sticking it out without me for two solid years. These have been the longest years in my life and I hope you will all find comfort in my achievements. Thank you for your support and encouragement.

Many thanks and appreciation goes to my supervisor, Professor Heinz Rüter, for encouraging me when the chips were down. Without your guidance this thesis would not have come to fruition.

To Solomon and Chiedza, thank you guys for the free tutorials that you gave me during my hours of need. I would like to thank all the people who assisted me in one way or the other during the preparation of this thesis. Your assistance nudged me a step closer towards the end.

## **DEDICATION**

**I dedicate this thesis to my grandmother, *Ambuya Zvakanaka*, she has always been my tower of strength. Ndinokudai vaChihera.**

# TABLE OF CONTENTS

Declaration .....	i
Abstract .....	ii
Acknowledgements .....	iii
Dedication .....	iv
1 BACKGROUND .....	1
1.1 Introduction .....	1
1.2 Problem Definition .....	2
1.3 Scope of Research .....	4
1.4 Research Objectives .....	5
1.5 Case Study Area .....	6
1.6 Research Approach .....	8
1.6.1 Literature Review .....	8
1.6.2 Design of conceptual LRBIP-EIS data model .....	8
1.6.3 Prototype development and implementation .....	8
1.7 Proposed Structure of Thesis .....	8
1.7.1 Background .....	8
1.7.2 Evolution of EIS in Africa .....	8
1.7.3 Software Development Methodology .....	9
1.7.4 Data Modelling .....	9
1.7.5 LRBIP-EIS User Needs Assessment .....	9
1.7.6 LRBIP-EIS System Design .....	9
1.7.7 LRBIP-EIS Prototype Development and Implementation .....	9
1.7.8 Evaluation and Recommendations .....	9
2 EVOLUTION OF ENVIRONMENTAL INFORMATION SYSTEMS IN AFRICA .....	10
2.1 Nolan's Model .....	10
2.1.1 Innovation .....	10
2.1.2 Contagion .....	10

2.1.3	Control .....	11
2.1.4	Integration .....	11
2.2	State of Environmental Information Systems Development in Africa .....	11
2.3	Problems Encountered during EIS Evolution .....	14
2.3.1	Technical Constraints .....	14
2.3.2	Institutional Barriers .....	15
2.3.3	Human Resources .....	15
2.4	Concluding Remarks .....	16
3	SOFTWARE DEVELOPMENT METHODOLOGY .....	18
3.1	Feasibility Study .....	19
3.2	System Definition .....	19
3.3	User Needs Assessment Techniques .....	21
3.3.1	Procedure for carrying out user needs assessment .....	21
3.3.2	Designing the user needs collection guide .....	22
3.3.3	Conducting Interviews .....	22
3.3.4	Analysing the interview results .....	23
3.3.4.1	Users' requirements .....	23
3.4	Database Design .....	24
3.4.1	System Configuration .....	24
3.5	Application Design .....	24
3.6	Prototyping and Implementation .....	25
3.7	Operational Maintenance .....	25
4	DATA MODELLING .....	26
4.1	Data Modelling .....	26
4.1.1	Conceptual Model .....	26
4.1.2	Logical Model .....	27
4.1.3	Physical Model .....	27
4.2	Database Structures .....	27
4.2.1	Hierarchical Data Model .....	28
4.2.1.1	Properties .....	28
4.2.1.2	Limitations .....	29

4.2.2	Network Data Model .....	29
4.2.2.1	Limitations of the network data model .....	29
4.2.3	Relational Data Model .....	29
4.2.3.1	Properties of relations .....	30
4.2.3.2	Limitations of the relational data model .....	31
4.2.4	Object-oriented data model .....	32
4.2.4.1	Limitations of the object-oriented data model .....	33
4.3	Entity-Relationship Modelling .....	33
4.3.1	Relationships .....	34
4.3.2	Cardinality of a relationship .....	34
4.3.3	Participation Constraints .....	36
4.3.4	Generalisation/Specialisation .....	36
4.4	Connection Traps .....	36
4.4.1	Fan Traps .....	36
4.4.2	Chasm traps .....	37
4.5	Normalization .....	38
4.5.1	Functional dependency .....	38
4.5.2	First Normal Form (1NF) .....	39
4.5.3	Second Normal Form (2NF) .....	39
4.5.4	Third Normal Form (3NF) .....	40
4.5.5	Boyce Codd Normal Form (BCNF) .....	41
5	LRBIP-EIS USER NEEDS ASSESSMENT .....	43
5.1	Components of an Environmental Information System .....	45
5.2	Fisheries .....	46
5.2.1	Fisheries information requirements .....	48
5.3	Apiculture .....	49
5.3.1	Beekeeping information requirements .....	50
5.4	Wildlife .....	51
5.4.1	Wildlife information requirements .....	52
5.5	Agriculture .....	53
5.5.1	Agriculture information requirements .....	54
5.6	Forestry .....	54

5.6.1	Forestry information requirements .....	55
5.7	Mining .....	55
5.7.1	Mining information requirements .....	55
5.8	Geodatabase activities .....	56
5.9	Current Computer System Configuration .....	56
5.10	Cooperation with other agencies/information systems .....	56
5.11	Conclusion .....	57
<b>6</b>	<b>LRBIP-EIS SYSTEM DESIGN .....</b>	<b>59</b>
6.1	Modelling tools .....	59
6.2	LRBIP-EIS conceptual data modelling .....	59
6.2.1	Fisheries conceptual data model .....	60
6.2.2	Beekeeping conceptual data model .....	62
6.2.3	Wildlife conceptual data model .....	64
6.3	LRBIP-EIS logical data modelling .....	65
6.3.1	Fisheries logical data model .....	67
6.3.2	Beekeeping logical data model .....	69
6.3.3	Wildlife logical data model .....	70
6.3.4	Validation of relations .....	72
6.4	Global logical data model .....	72
6.5	Physical data model .....	72
<b>7</b>	<b>LRBIP-EIS PROTOTYPE DEVELOPMENT AND IMPLEMENTATION .....</b>	<b>74</b>
7.1	Introduction .....	74
7.2	Prototype Scheme .....	74
7.3	Fisheries Test Database .....	75
7.3.1	Non-spatial database .....	76
7.3.2	Spatial database .....	79
7.4	Fisheries Prototype User Documentation .....	80
7.4.1	Setting up LRBIP-EIS prototype .....	80
7.4.2	Navigating through the system .....	80
7.4.2.1	Logging in .....	80
7.4.3	Data capture and update .....	82

7.4.4	Data Querying .....	84
7.4.5	Report Generation .....	84
7.4.6	Data Security .....	86
7.4.7	Limitations of the prototype .....	86
8	EVALUATION AND RECOMMENDATIONS .....	88
8.1	Evaluation .....	88
8.2	Recommendations .....	89
	REFERENCES .....	91
	APPENDIX A: USER NEEDS ASSESSMENT GUIDE .....	98
	APPENDIX B: SAMPLE DATABASE TABLES AND DATA DICTIONARIES .....	102
	APPENDIX C: LRBIP-EIS SET UP CD WITH SOURCE CODE .....	104

# 1 BACKGROUND

## 1.1 Introduction

Mankind must attempt to find the delicate balance between its developmental needs on the one hand and the preservation of the status quo, or preferably the reversal of environmental damage, on the other. While it is highly unlikely, if not entirely unrealistic to assume, that environmental management models will provide for full sustainability, every effort must be made to minimise negative human impact on the environment. It is well known that increased population and socio-economic developments, such as urbanisation, apply tremendous pressure on agriculture, biodiversity, climate, vegetation, wildlife, and water resources, among others. Man's exploitation of these resources inevitably leads to environmental degradation in the form of soil erosion, deforestation, increased urbanisation and irregular settlements (EIS-Uganda, 2002). Conservation strategies should prevent depletion of resources; minimize ecological imbalances and maintain bio-diversity. This calls upon government authorities to implement sound resource management practices that are concerned with the optimum allocation and exploitation of resources to ensure sustainable coexistence between stakeholders with the environment. Sustainable development can be viewed as the dynamic equilibrium between inputs and outputs of the development process which aims at continued utilization of available resources without endangering the environment. Agenda 21, established at the 1992 United Nations Conference on the Environment and Development (UNCED), in Rio de Janeiro, Brazil, addresses the development of societies and economies by focusing on the conservation and preservation of the environment and natural resources.

Environmental information systems are concerned with the collection, analysis and representation of data about the environment. This data comes from a wide array of sources such as meteorology, forestry and pastures, agriculture and fisheries, water resources, health, industry, transportation and urban infrastructure, mining, wildlife, and satellite imagery (Kalensky *et al*, 1998). The environmental data is spatially referenced

using geodetic control networks or aerial photogrammetry and remote sensing techniques. This permits the integration of sectoral data types to create aggregated environmental information systems (EIS) that can be utilized in policy development and decision-making.

Environmental Information Systems are based on a variety of advanced technologies such as Geographic Information Systems (GIS), Global Positioning Systems (GPS), remote sensing and image analysis systems as well as geospatial information networks; for the acquisition, processing, storage and dissemination of geospatial data (Kalensky *et al*, 1998). Geographic Information Systems and their associated databases are the fundamental components of EIS because of their capability to combine spatially referenced information and other relevant data into computer-based integrated information systems, which can be utilised by decision-makers, planners, scientists and experts, to carry out spatial analyses, provide geo-statistical data and display and report the results. (Ruther, undated). An Environmental Information System EIS is not just a technological tool, but also includes strategies, procedures and institutional frameworks, together with data management tools, that ensure access to environmentally relevant data and allow for their analyses (World Bank Group Report, 1999). It provides a powerful management tool and a heuristic approach to the decision-making process and development planning at all levels of governance by providing environmental practitioners with the appropriate kind of information at local, district, national and regional levels. This enables managers to identify and quantify specific environmental disciplines and to determine their optimum utilization because decision-making without reliable and up-to-date information in a form suitable for the task at hand is nothing more than guesswork (Humphries in Dahlberg *et al*, 1989).

## **1.2 Problem Definition**

Designing and developing integrated GIS-based environmental management information system that provide access to environmentally relevant data pose great challenges to developers. For example, GIS applications in environmental management have

predominantly taken a sectoral approach, focussing on particular areas such as fisheries, wildlife or forestry (MacDevette *et al*, 1995). These developments have tended to produce incompatible and inconsistent heterogeneous databases that require extensive effort to identify, acquire, comprehend and integrate (Slagle in Michener *et al*, 1994) due to the diversity of environmental data. This is a result of a number of actors being involved in the management of common environmental sites. Humphries (in Dahlberg *et al*, 1989) argues that integrated spatial information systems are urgent requisites lest it may be too late to unravel the computerised mess of largely departmental *environmental information systems*.

The complexity of the natural environment makes it almost impossible for central government to effectively conserve and manage the natural resources base single-handedly. In Tanzania, the National Environmental Policy (1998) and the Local Government Reform Act (1998) devolved governance authority to Regional, District and Village Councils to plan and manage local environmental policies and regulations (RFIS, 2001).

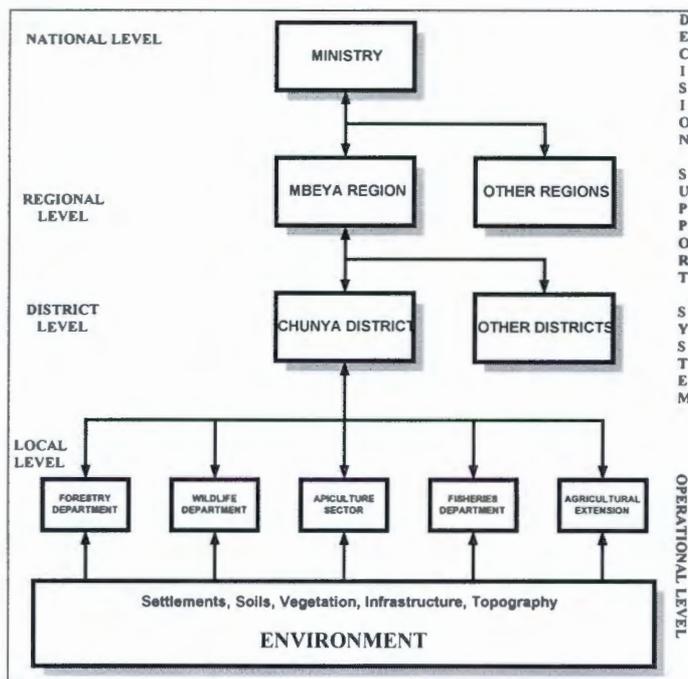


Figure 1.1 Environmental management framework (adapted from Bell *et al*, 1995)

These initiatives involve empowering local communities to participate in the decision-making process with regards to the resources found within their environs. Figure 1.1 (above) shows the management framework to implement the environmental policy in Tanzania. Central government is involved more with policy formulation (Bell *et al*, 1995) while the local communities assume direct responsibility for the exploitation of natural resources in a sustainable way so as to protect and conserve the environment (Chiwandamira *et al*, 1999). The country is divided into regions that are made up of districts. Each district has environmental departments such as forestry, wildlife and fisheries that work with the local communities and provide expert advice on the exploitation of natural resources at the local level.

Community-based conservation systems need close monitoring to prevent the over-exploitation of ecologically sensitive areas that may disrupt the eco-balance. This creates a major demand for environmental information and appropriate tools to manage it at the appropriate levels of decision-making, such as policy-making at central government level or quota determination and allocation at the local community level. The lack of computerised data collection, management and dissemination systems within government and individual sectors (*such as fisheries, apiculture, agriculture, mining, forestry and wildlife*), responsible for environmental management results in fragmented environmental information systems. This is compounded by poor data modelling, capture, storage and analysis structures adopted in designing and implementing some of these information systems.

### **1.3 Scope of Research**

Environmental Information Systems are based on the “distributed model” (NEMC, 1998), in which each individual sector has custody of data relevant to its applications but is encouraged to develop and host interoperable environmental data sets. The sectoral data should be in a format that allows it to be used by other environmental stakeholders who may be operating using different software configurations. Formal and/or informal data sharing arrangements impart a structured process to the design, development and

management of environmental data. It is this structure, if well designed, that creates a cost-efficient approach to the multiple use of multisource data (Slagle in Michener et al, 1994). The success of the distributed model requires strong networking capacity, coordination and harmonisation of the different datasets for the production of integrated geo-spatial databases.

This research will focus on database development methodology in the design of a GIS-based integrated-environmental-information-system database following the concepts of the "distributed model". The relational database model will be used as the underlying database structure. A case study of Lake Rukwa Basin will be carried out and an integrated database will be developed incorporating the following environmental sectors; agriculture, wildlife management, forestry, fisheries, and apiculture. Prototyping will be confined to two or more components of the environmental information system due to time constraints and timely availability of data. The maintenance and review of the system is beyond the scope of this research but recommendations will be made to this effect.

#### **1.4 Research Objectives**

The aim of the research is to develop a GIS-based integrated environmental information system database for the management of natural resources in ecologically sensitive areas. The Lake Rukwa Basin Project in Tanzania; hereafter referred to as Lake Rukwa Basin Integrated Project (LRBIP), will be used as a case study.

The specific objectives of the research are:-

1. Review of the evolution of environmental information systems in Africa
2. Determine the information needs of the LRBIP
3. Design the conceptual, logical and physical data models for LRBIP database
4. Develop and implement a prototype of an integrated GIS-based LRBIP database

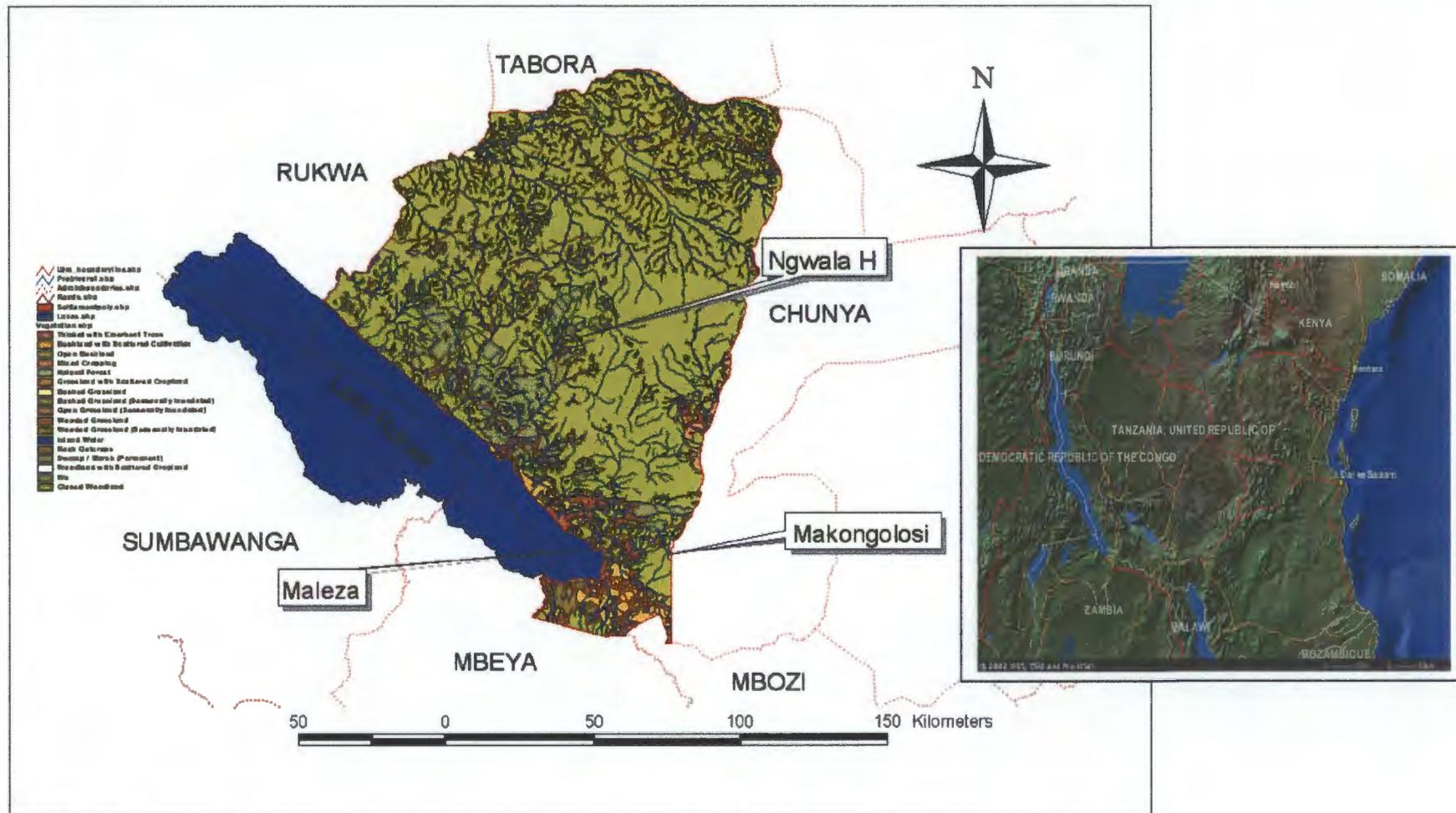
## 1.5 Case Study Area

Lake Rukwa Basin is located in Mbeya region on the south-western part of Tanzania and is part of the Rift Valley with Lake Tanganyika on the northwest and Lake Malawi on the southwest. The area of Lake Rukwa, which strides the regions of Mbeya and Rukwa is isolated compared to the main economic development axis of Tanzania. This area is classified as an area of low productive potential, except for some restricted zones that are suitable for corn, tobacco and cotton production (Davenport, 2000). According to the same document, the districts of Chunya and Mbeya Rural have a rich endowment of natural resources with a wealth of fauna and flora species. The ecological balance of this region is under threat from increased human population demands on the natural resource base.

The Lake Rukwa Basin Integrated Project (LRBIP) is a result of the bilateral agreement signed on 27 February 1991 between the governments of Italy and Tanzania. The main objective is the sustainable utilisation of the natural resources for the benefit of the local communities living in the Lake Rukwa Basin and its environs. The project began in May 1994 and is being funded by the Italian organization Centro Internazionale Crocevia (CIC), the Government of Tanzania and the Italian development Cooperation.

According to R  ther (2001), the case study area covers approximately eighteen thousand (18 200 km<sup>2</sup>) square kilometres and has a diverse range of natural resources and human activities. It is made up of the entire part of Chunya district west of the Itigi-Makongolosi road bounded by (Figure 1.2 overleaf):

- The Northern border of the Chunya district from its intersection with the eastern lakeshore to its intersection with the Itigi-Makongolosi road
- The Itigi-Makongolosi road from its intersection with the Chunya district border to Makongolosi, including focus villages
- A line due south (meridian) from Makongolosi to the Chunya district border



**Figure 1.2 Lake Rukwa Basin (Case Study Area)**

- The southern border of the Chunya district from its intersection with Makongolosi meridian to its intersection with the lakeshore
- The lakeshore from its intersection with the southern Chunya border to the intersection of the northern Chunya border with the lakeshore, including the portion of Lake Rukwa which falls into the Chunya district.

### **1.6 Research Approach**

The literature review covered a wide range of topics so as to satisfy the objectives outlined in section 1.4. The literature review included the following areas:

- Review of environmental information systems, GIS-based natural resources information systems and the role of GIS in environmental management
- Review concepts of database design, prototyping and implementation
- Review of the environmental activities in the Lake Rukwa Basin
- Study of the software (such as Microsoft Access, ArcView GIS, Map Objects and Visual Basic) to be used for developing the prototype

### **1.7 Structure of Thesis**

The thesis has been divided into eight chapters. Chapter One is an introductory chapter which gives an overview of environmental information systems, defines the research problem and states the scope of the research work. It gives a brief overview of the case study area. Chapter Two outlines the evolution of environmental information systems (EIS) in Africa in the context of Nolan's model of IT adoption and organisational learning. It highlights some of the problems that have beset EIS initiatives and presents various ways of mitigating these problems.

Most successful database applications are based on the Software Development Methodology. Chapter Three reviews this methodology by analysing the stages of the database application lifecycle. A user requirements collection guide was designed for use during user needs assessment. Chapter Four focuses on data modelling techniques, which are essential in the design of any information system. It also reviews the common

database structures, such as hierarchical, network, relational and object-oriented. A review of entity-relationship diagrams and normalisation is also made in this chapter.

Chapter Five details the results of the user-needs assessment. The assessment was carried out using techniques such as interviews, questionnaires and reviews of completed or ongoing environmental management information systems. Chapter Six uses these results to develop the data models for individual sectors (fisheries, apiculture and wildlife). The separate data models are then integrated into a global logical data model by eliminating areas of duplication between the sectors and using the areas of common interest to bridge the data models. The integrated data model was used for developing a prototype of the LRBIP-EIS database application. Chapter Seven outlines the prototyping procedures used for the LRBIP-EIS database. It also explains the graphical user interface (GUI) development and describes how the different software components were integrated into the prototype application.

The final chapter of the thesis presents an evaluation of the research by looking at the objectives the author set out to achieve. Chapter Eight also outlines some recommendations based on the assessment of the research. These include areas identified for possible future research activities.

## **2 EVOLUTION OF ENVIRONMENTAL INFORMATION SYSTEMS IN AFRICA**

As outlined in Chapter 1.1, Environmental Information Systems (EIS) can be viewed as extensions of Geographic Information Systems (GIS), that synthesize environmental data to produce information required for solving complex planning and environmental management problems. The establishment of environmental information systems, particularly in African countries, has followed the traditional approach to information technology adoption as predicted by Nolan. Nolan's model outlines the evolution of information systems in four developmental stages: innovation, contagion, control, and integration (Yeates *et al.*, 1994), which according to Lyytinen (1991), has its own distinct characteristics.

### **2.1 Nolan's Model**

#### **2.1.1 Innovation**

During the innovation stage, computers are used to satisfy basic needs of very few enthusiastic individuals within an organisation. As the use of computing technology gathers momentum, computerisation problems will inevitably arise. These problems are further compounded by the fact that typically minimal planning is done before the establishment of computing facilities. At this stage management is typically least concerned about these problems since they are not the major focus of the organisation's activities.

#### **2.1.2 Contagion**

Successful implementation of information technology (IT) by few individuals often triggers a rapid increase in the use of computing within the organisation. Management will begin to realise the great potential of this technology while most users' expectations increase. This results in parallel developments of computing applications leading to duplication of computing facilities and the adoption of different specifications for hardware and software.

### **2.1.3 Control**

At this stage, most organisations establish computing departments to coordinate the various computing initiatives in their bid to plan, control and formalise the growth of the technology. The position of IT management in the organisation is well acknowledged, often leading to controlled standardisation of hardware and software configurations. The information system planning is given high priorities and management controls the costs. Data processing becomes centralised, creating a single information system for the whole organisation.

### **2.1.4 Integration**

Information systems have the tendency to grow in leaps and bounds, reaching unmanageable proportions within short periods of time. As the system grows, control structures are re-evaluated, sometimes leading to the decentralisation of application development. Use and application development is rationalised and coordinated. Planning is widely accepted and any centralisation or decentralisation of computing resources and applications is controlled through business strategies.

## **2.2 State of Environmental Information Systems Development in Africa**

Environmental Information Systems development in Africa is going through the different stages of the Nolan's model with different countries at different stages of development. The continent as a whole is at the control stage and efforts are being put towards the establishment of national EIS frameworks with the view of expanding these to regional and eventually continental frameworks.

Technological advances in the field of remote sensing from the early 1970s to the early 1980s, marked the inception of environmental information systems. These developments, such as the launch of remote sensing satellites (LANDSAT, SPOT, etc), provided large amounts of environmental data that could be used for analysis and management. Few enthusiastic individuals championed the establishment of environmental information systems at departmental levels of various institutions concerned with environmental management. The majority of EIS related activities during this decade were exploratory

or experimental in nature and confined to specific sectors with few, if any, linkages between sector efforts. Most of the environmental information systems were created to support donor-funded (UNEP, UNDP, FAO, etc) projects such as environment support projects and natural resources management projects and were, as a result, supply-driven and, project as well as data-oriented. Pockets of expertise (information communities) in the fields of GIS, remote sensing and database management systems (DBMS) technology developed in most African countries during this era.

During the 1980s, more and more institutions became aware of the need to establish environmental information systems due to increased pressure on natural resources from rising population levels as well as natural disasters such as floods and drought. This period saw a phenomenal growth in the number of actors involved in EIS construction. Duplication of data and resources during this stage was inevitable due to the legacy of sectoral environmental management policies. These policies delegated different government ministries to manage different sectors of the same environment. In many countries, for example, the departments of forestry and wildlife are separate institutions run by different ministries. The result was a multitude of EIS groups operating as an unruly collection of factions pulling in different directions, each driven by its own valid objectives (Prévost and Gilruth, 1999). In Zimbabwe, for instance, the Integrated Resources Information System (IRIS), Vegetation Resources Information System (VEGRIS) and Agricultural Land Evaluation Information System (ALEIS) initiatives had remote sensing activities but were funded by different donors and were completely uncoordinated. The national institute for remote sensing was not involved in the development of these information systems (EIS-Zimbabwe, 2002). This state of EIS development posed a crisis because the lack of coordination retarded sustainable development and promoted conflict between the different sectors. Discussions on EIS policies to try and control the uncoordinated growth of the environmental information systems were mooted.

Control of the development of EIS initiatives began in the mid-1980s, as a result of the recognition that environmental information was a distinct cross-sectoral issue through the

adoption of National Plans to Combat Desertification and National Conservation Strategies. The same realisation came out of the National Environmental Action Plans (NEAP) processes initiated in the late 1980s, which emphasised the need for shared solutions and integrated data products. The 1992 UN Conference on Environment and Development (UNCED) in Rio de Janeiro called for the establishment of information systems that would improve access to information with environmental relevance and make it available as a basis for decision-making. This challenged the environmental information communities to recognise their mutual interest and work towards a greater synergy of their respective efforts. EIS initiatives were propelled from a supply-driven to a demand-driven orientation. The underlying principle is that EIS should serve a clearly specified management need, and that data should not be collected unless an end use is defined (Prévost and Gilruth, 1999). Some African governments have responded to this challenge by formulating holistic environmental management policies. The World Bank has been funding the Environmental Information Systems for Sub-Saharan Africa (EIS-SSA) Program (now EIS-AFRICA) since the early 1990s, to promote the implementation of effective environmental information systems. The Program supports African countries as they assess their priority needs in terms of environment and land information systems, and analyse the technical, institutional, legal and economic issues hampering their possibilities of meeting these needs (EIS-SSA document, 2001).

Environmental management is effective if it is integrated into the decision-making process at all levels. The complex nature of the environment makes the distributed model ideal for the establishment of EIS. Here, data sets are constructed and hosted by institutions with the appropriate statutory mandates under the control and supervision of capable data custodians. The recognition and support of data custodians is crucial for EIS development and they should be encouraged to develop a culture of data sharing and to have policies that minimize duplication of resources. Countries such as Uganda, Ghana, Zambia, Eritrea and Tanzania have launched initiatives to establish national environmental information networks (EIN) (UNEP document, 2002). These frameworks aim at minimizing the institutional and technical constraints to EIS development by providing horizontal and vertical structures for sharing data within the environmental

information communities. It is relatively easy to coordinate members of the EIN when their disparate data sets are being aggregated to fulfill high-level management objectives such as State of the Environment Reports (SOERs). The Ghana - Country at A Glance (G-CAG) database was created by generalizing and synthesizing data from several custodians. It is a synoptic, inter-operable, and user-friendly geographical database designed to assist in national-level environmental management and planning. The aim was to construct a versatile and inter-operable geographic database directed towards decision-makers and similar persons who need to have an overview over the country for large area planning purposes. Another aim with the data sets in the Ghana-CAG is to serve as an introduction to the detailed data sets that are available at the custodian organizations (EIS News, 1999). EIS initiatives such as the Peace Parks and Lake Tanganyika Biodiversity Program have transcended national boundaries and are being used to generate valuable environmental information on shared resources.

### **2.3 Problems Encountered during EIS Evolution**

The development of EIS has been fraught with numerous problems ranging from institutional barriers to technical constraints as well as limited human resources capacities. These inhibiting factors have been well documented in World Bank reports on cases in Best Practices of EIS for countries such as Zimbabwe, Ghana, Mozambique, Uganda and Senegal. These problems are consistent with predictions of Nolan's model for IT adoption during the early stages.

#### **2.3.1 Technical Constraints**

Most African governments operate under stringent financial constraints and usually fund projects that have immediate political and socio-economic gains. Funding for EIS projects with long-term gains and that required huge capital expenditure, was therefore, not generally readily available from central governments. Donor-funded projects in this sector became the norm. This resulted in a number of problems including the proliferation of incompatible hardware and software configurations in implementing agencies. The lack of project coordination resulted in the adoption of different database development standards on different projects. The wide spectrum of data sources resulted

in data integration problems caused by general lack of standards for map projections and coordinate systems, naming conventions, and accuracy standards. This posed, and still poses, severe data harmonisation difficulties when the disparate sources have to be integrated into a single format. As an added complication it became obvious that the electronic communications facilities in most African countries could not meet the demands of the distributed nature of EIS data and facilities.

### **2.3.2 Institutional Barriers**

Institutional barriers arise from the legacy environmental management frameworks, which were largely sectoral. Most organisations lacked a coordinated participatory approach, which is crucial for the success of EIS implementation. This could be attributed to the resistance to central coordination as some of the environmental practitioners considered this as a threat to their autonomy. Although the effective operation of GIS-based EIS depends on the availability of historical and current geospatial data collected by a number of government agencies, Kalensky *et al*, 1998; notes that most agencies consider their databases as their exclusive property, which they are not prepared to share with anybody. In most countries, there was no clear policy on environmental management. Organisations carried out their functions independent of each other and lacked vertical and horizontal networking to improve data access and sharing. The legal mechanisms for inter-sectoral information exchange are virtually non-existent in most countries. The fact that copyrights to environmental data are not clearly stipulated is of major concern in most organisations in terms of data dissemination.

### **2.3.3 Human Resources**

During the infancy of EIS implementation, most projects depended on donor-funded expatriates due to the lack of sufficiently trained local expertise. The expatriates were usually employed on a contract basis and most projects were not sustainable as soon as the experts left. Local personnel, trained as part of capacity building in most projects were often offered higher managerial posts that removed them from the technical aspects of these projects. This was compounded by the failure of African academic institutions to produce enough personnel in the field of GIS and environmental management (Rüther,

2001). The training often concentrated on environmental GIS technology rather than on environmental information management.

#### **2.4 Concluding Remarks**

Environmental Information Systems provide an integrated source from which decision-makers and planners can draw authentic and viable information on environmentally related issues. WLIP (1991) advocates the development of a decentralised confederation of systems based on common standards, where those with land information responsibilities (data custodians) would continue to collect, maintain and keep custody of that information. The EIS community needs to continue to develop new applications that encourage members to share information for their mutual benefit. Uganda is a good example of a country that has a decentralised environmental information network, where the environmental management policy is enshrined in that country's 1995 Constitution and a National Environmental Action Plan is implemented at district and local council levels under the auspices of a coordinating body, the National Environmental Management Authority (EIS-Uganda, 2002). The distributed approach promotes greater flexibility and discretion for participating institutions in developing their own environmental information systems while encouraging inter-disciplinary cooperation and networking. Formal and/or informal data sharing arrangements, encompassing standards and conventions, should be in place to promote integration and interoperability of environmental information.

Environmental information systems should be demand-driven i.e. based on thorough needs and institutional analysis with focus on information requirements. On a strategic level, database initiatives in various projects should be coordinated as much as possible to share identified needs and design ideas. This would improve the usage and compatibility of systems, allowing improved sharing and comparison of information among members of the EIS community. A common data infrastructure is the backbone for the integration of distributed information systems.

EIS practitioners should consult National Mapping Organisations (NMOs) prior to and during EIS implementation, so that EIS activities are closely linked to National Spatial Data Infrastructure initiatives. In cases where these are in place, this will greatly improve the harmonization of data architecture and reducing duplication to minimum levels. Smith (1999) notes that GIS organisations and users will make significant cost savings through reduced duplication or recapture of spatial data and the use of metadata i.e. data about data, to locate the available datasets.

The training of EIS practitioners should combine technology training and environmental information management (Ruther, 2001). This should play a significant role in the formalisation of the development of EIS in Africa. Collaboration between the EIS community and academic institutions should encourage universities to undertake research into some of the problem areas such as environmental database design and implementation. Geomatics institutions in countries such as Ghana, Democratic Republic of Congo, Kenya, Nigeria, South Africa, Tanzania, Zambia and Zimbabwe should be encouraged to incorporate EIS into their curricula and to be research partners with implementing agencies and organisations.

EIS initiatives should embrace new technologies such as the World Wide Web (WWW), Internet and multi-media, so that decision-makers and planners have ready access to environmental data. EIS should take advantage of developments in the electronic communications sector where these are available or stimulate the growth of this sector in areas where it is lagging behind.

The full potential of EIS is, with a few exceptions, far from being realised in Africa due to the prevailing state of technological, political and economical development in most countries. The development of environmental information systems should be encouraged as they provide the decision-makers with accurate information on which to make informed decisions about the sustainable utilization of available natural resources.



The stages are not strictly sequential, but involve some amount of repetition of previous stages through feed-back loops (Connolly et al, 2002). For example, problems encountered during database design may necessitate additional user requirements collection and analysis. For brevity, Figure 3.1 shows only some of the main feed-back loops between the stages. This chapter presents a brief description of each of these stages. User needs assessment techniques are discussed in detail while database design, and prototyping issues will be dealt with in detail in Chapters Six and Seven respectively.

### **3.1 Feasibility Study**

Feasibility study is conducted after the identifying in very broad terms, a problem(s) warranting the introduction or improvement of an information system. The aim of this study is to determine the resources (i.e. financial, human and technical) required to solve the problem and the enterprise's capability to meet the demands of these resources. A steering team that has decision-making powers within an enterprise usually carries out this study. The team should have representatives from management (e.g. directors, managers), participating sectors (data custodians, users at the factory floor), and information technology (IT) departments (Wiggins *et al*, 1991). This coordinated approach will promote the creation of a shared information system with clearly defined mission statement and objectives. Such a system should minimize data redundancy in maintaining separate data sets (e.g. map records) in several departments which should improve the provision of more consistent information. At the end of the feasibility study, the steering team should be able to define the system boundaries and make recommendations in terms of identifying and allocating financial, human and technical resources.

### **3.2 System Definition**

Boundaries of the proposed information system are defined by using data flow diagrams (DFD) that show the flow of information between the current and potential components of that system (see Figure 3.2 below for an example). The components include both the internal and external entities that interact with the proposed system (Hawryszkiewicz, 1991). Boundary definition makes it relatively easy to identify all the departments/sectors

that are major users of geo-spatial information within the organization that should be candidates for requirements collection. The systems analysts can then carry out detailed user needs analysis according to individual user or department job role or application area (Connolly *et al*, 2002).

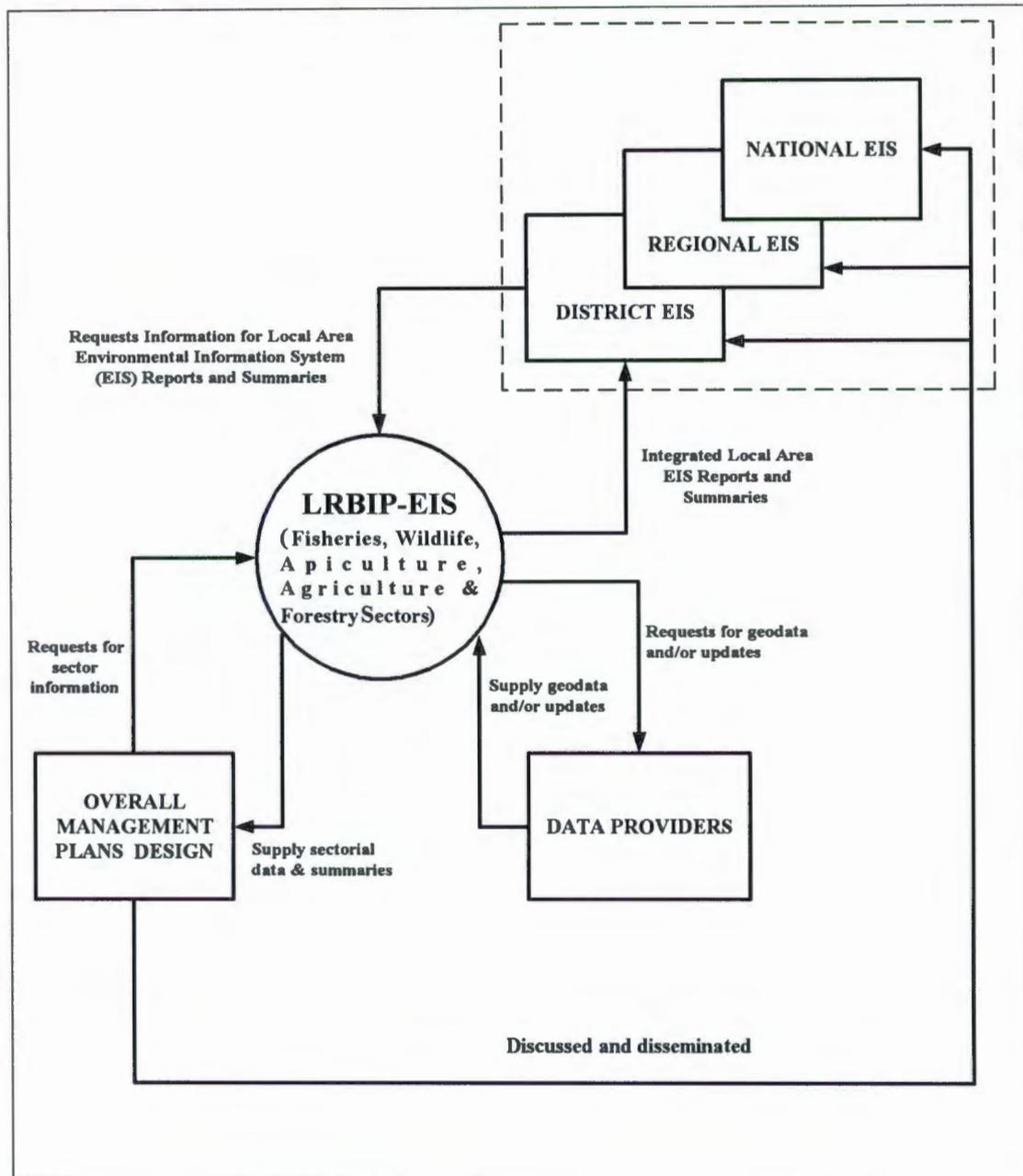


Figure 3.2 Data flow diagram

### **3.3 User Needs Assessment Techniques**

User needs assessment can be carried out using the centralized approach, the view integration approach or a combination of the two, depending on the degree of overlap of the different sectors' (users) requirements within of an enterprise. In the centralized approach, all stakeholders come together and produce a single set of requirements for the new system. This approach is preferred when there is a significant overlap in requirements among users. The view integration approach is preferred when there are significant differences between sector user needs. This approach identifies the user requirements for each sector and uses them to build separate conceptual data models that represent each user's perspective of the new system. In enterprises that are more diversified, a combination of the two approaches can be used with the centralized approach being employed for those sectors that have a lot in common while the view integration approach can be used for those sectors with distinct needs. The outcome of the users requirement assessment will determine whether a single (centralized) or shared (distributed) information system is developed for the enterprise. In a distributed information system, interoperability of data sets becomes a critical factor because more than one system will be necessary to serve the information needs of the enterprise. In such cases, if the distributed systems are not networked, then they should be designed to allow data such as the same mapped information to be shared between several GIS systems that have special analytic capabilities (Wiggins et al, 1991).

#### **3.3.1 Procedure for carrying out user needs assessment**

The procedure involves a number of fact-finding techniques depending on the state of the information system of the enterprise under investigation. For existing systems, the procedure involves interviewing the current and potential users of the system, reviewing of any detailed systems documents, conducting literature and industry researches on similar operational systems, observing the enterprise in operation and the use of questionnaires. Where completely new systems are proposed, the procedure relies heavily on interviews with particular emphasis on users' expectations of the new system (Hawryszkiewicz, 1991).

The information gathering is carried out in a top-down manner with a high level system developed first from the management perspective. Goodwin *et al*, 1993, encourages GIS and business analysts to understand what the components of the enterprise aim to achieve and how they currently operate, and what new ways the users would like to employ if given the capabilities of GIS technology. The new system must take into account the operational requirements of the everyday users (Hawryszkiewicz, 1991). Guidelines should be drawn up for the information gathering process to succeed.

### **3.3.2 Designing the user needs collection guide**

For any user needs assessment to be successful, it must be carried out in a consistent and structured manner. A requirements collection guide that outlines what basic information should be collected from each respondent will provide a consistent framework for determining the needs of the various departments and sections within the organization (See Appendix A for an example). This guide can be used by interviewers as a checklist during interviews with the various users.

### **3.3.3 Conducting interviews**

Open-ended, personal interviews are the appropriate vehicle for user needs assessments because they allow the interviewer to follow up lines of questions about particular application and analysis needs, and to get detailed answers about a department's projected use of maps and databases (Wiggins *et al*, 1991). Personal interviews allow the interviewer to collect necessary details and/or requests and obtains copies of supporting documents directly during the interview process (GIS Development Guidelines, 2001). Individual interviews are preferable to group ones because some questions may prove sensitive and also, extremely vocal individuals tend to dominate group interviews (leading to bias). Furthermore, group meetings easily lose focus on specific GIS applications and therefore do not provide detailed information needed to adequately describe GIS applications (GIS Development Guidelines, 2001). The interview sessions should not be unnecessarily long. Wiggins *et al*, 1991, suggest limiting the sessions to between one and two hours. For large and complex departments, at least one interview should be conducted in each section. During the course of the interviews, it may be

realized that some essential people had been omitted from the interview list; these people must be added and interviewed for their input.

Interviews can be conducted by professional consulting firms, university departments, or by in-house teams. Professional consulting firms and university departments have the advantage that they are very current with regards to changes in information technology in terms of both hardware and software and their recommendations are unbiased. In-house teams may prove very educational to those participating in the information gathering process but there is a high probability of heavily biased assessment due to individual or departmental inclination and more importantly the issue of fear of becoming redundant.

### **3.3.4 Analysing the interview results**

The interviewer should write down a summary of each interview as soon as possible, while his or her notes are still fresh. Follow up interviews should be carried out as and when necessary. The summaries form the basis for users' requirements, estimated costs and benefits, hardware and software configurations (systems requirements) and implementation plans for the proposed information system (GIS Development Guidelines, 2001).

#### **3.3.4.1 Users' requirements**

A sector-by-sector summary should itemize existing conditions (maps, databases, hardware and software use), current decision-making functions and data used to make those decisions, and the principal sources and destinations of shared data (Wiggins *et al*, 1991). The information flow should be summarized using matrices and flow diagrams for both spatial and non-spatial data sets. Typical matrices have rows that correspond to particular maps or attribute data, and columns that correspond to the users/producers of the maps and data (Wiggins *et al*, 1991). Flow diagrams are useful to illustrate where maps or databases are originated, which departments then use the information, and finally, where changes and maintenance occur.

The analysis of the user requirements should identify existing problems such as information bottlenecks, weak and missing data. Particular note should be made of problems with getting information on a timely manner, the scheduling of corrections and updates, and redundant data entry. Users expectations of future capabilities of the proposed system should be summarized for the organization as a whole (Wiggins *et al*, 1991).

### **3.4 Database Design**

The results of the user needs assessment are used to design the database application. The database design process is characterized by the conceptual design, logical design and physical design stages. The conceptual design stage includes identification of important entities, their relationships and their associated attributes. The logical design stage maps the conceptual model on to a logical model that is influenced by the data model of the underlying DBMS such as relational, or object-oriented. The physical design allows the designer to make decisions on how the database is to be implemented (this is closely tailored to a specific DBMS).

#### **3.4.1 System configuration**

Specifications for the system hardware and software configurations should be drawn up soon after the design of the conceptual data model. The user requirements and the availability of resources dictate the hardware and software configuration of the proposed information system. The software configuration depends on the type of database management system, performance, security, backup and recovery requirements of the proposed system. The hardware configuration may be wide area network (WAN) server-based systems with terminals, local area network (LAN) systems or stand-alone personal computers (PC) depending on the outcome of the user requirements analysis.

### **3.5 Application Design**

Application design involves the design of the user interface and the application programs that use and process the database to meet the functionality stated in the user needs specifications. Application design and database design are parallel activities because the

database exists to support the applications, and so there must be a flow of information between application design and database design (see Figure 3.1).

### **3.6 Prototyping and Implementation**

Information systems can only be developed over time and the implementation plans must be designed to take this into account. The plan might include proposed costs, phasing and scheduling (Wiggins *et al*, 1991), and the data conversion, database population and testing routines. Depending on the size of the organization, its existing automation, its growth rates, and its financial capabilities, a prototype (evolutionary prototype) may be developed to provide a working model of the proposed system. The prototype does not normally have all the required features or provide all the functionality of the final system. It is assessed by the users to check if the proposed system meets their requirements. Once the requirements are satisfied, the prototype will be developed further to become the working database application (Connolly *et al*, 2002).

### **3.7 Operational Maintenance**

The maintenance phase involves monitoring the performance of the system, and maintaining and upgrading the database application. If the performance falls below an acceptable level, tuning or reorganization of the database may be required.

## 4 DATA MODELLING

### 4.1 Data modelling

Data modelling is an abstraction of the real world that enables system designers, programmers and end-users to gain a common understanding of the nature of the data and how it will be used in an enterprise (Connolly *et al*, 2002). An enterprise is the organization implementing an information system. According to Laurini *et al*, 1992, there are three different levels of data modelling (conceptual, logical and physical) that provide the framework for the design of a computer-based information system (Figure 4.1 below). These levels are usually based on the principle of working from the whole to the part, the whole being the real world phenomena.

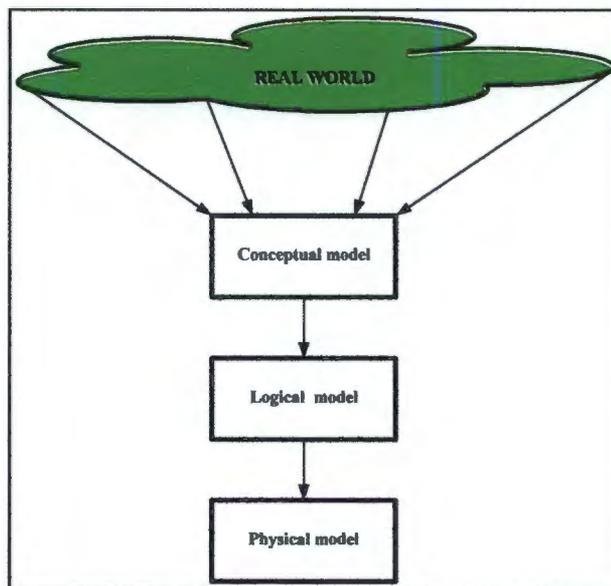


Figure 4.1 Data modelling levels (based on Laurini *et al*, 1992)

#### 4.1.1 Conceptual model

The conceptual model captures and represents the users' perspective of the information system (external view of the end-users) as perceived by the designer. During the conceptual modelling phase, the end-users define and describe the system requirements without taking into account the implementation aspects (Atzeni *et al*, 1999) such as

hardware and software requirements. Entity-relationship modelling is used to graphically represent the relationship between phenomena (Laurini *et al*, 1992). This provides an effective way of communicating with the end-users when identifying information requirements of the proposed system because system designers, programmers and end-users tend to view and interpret the system requirements in different ways. For large and complex systems, separate conceptual models are developed to accommodate individual users' expectations of the new system.

#### **4.1.2 Logical model**

The logical modelling transforms the conceptual data model(s) into a data model (such as relational (4.2.3), object-oriented (4.3.4), network (4.2.2) or hierarchical (4.2.1) of the implementing database structure but without the technical implementation details. In the case where separate conceptual models are developed but the user requirements specify a single (centralized) system, the resulting logical data models are merged together to create a global logical data model. For the relational data model, the results of logical modelling include a data dictionary, determination of relations, attributes, attribute domains, entity and referential constraints. Normalisation (4.5) is used to validate the logical model to ensure that it is free of anomalies that may result in duplication or redundancies in the database.

#### **4.1.3 Physical model**

The physical model depends on the specific database management software (e.g. MS Access, Oracle, Sybase etc) chosen for implementing the database application. It takes into account the criteria for the internal organization of the data in the particular database management system (Atzeni *et al*, 1999) and describes how the data is stored in the database e.g. tables, records, formats, indexes etc.

### **4.2 Database Structures**

Database structure refers to the data model according to which the data is physically organized in a database. Four generations of widely used database structures (hierarchical, network, relational, and object-oriented in that order) have evolved over the

years to meet the increasing demands of database applications. A brief synopsis of each of these data models is presented in the following sections.

#### 4.2.1 Hierarchical data model

The hierarchical model was one of the first popular database models and it reflected the nature of typical applications of 1960s (such as accounting systems, payroll systems and airline reservation systems).

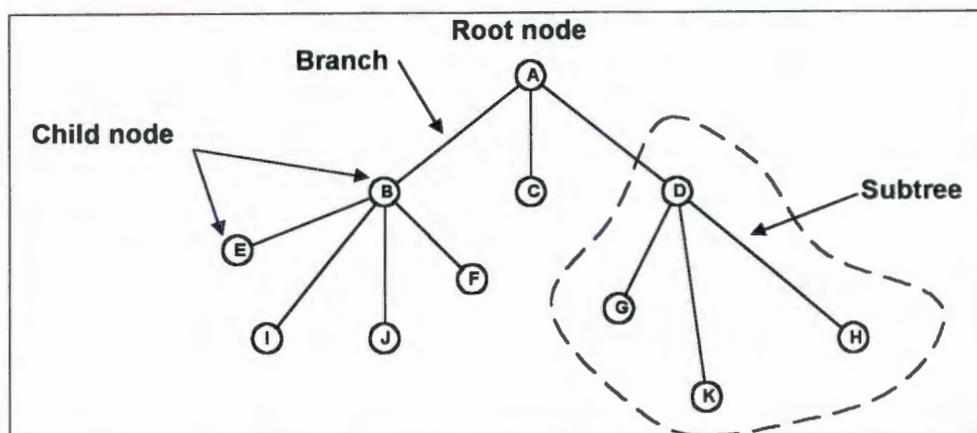


Figure 4.2 Hierarchical tree structure

##### 4.2.1.1 Properties of the hierarchical data model

The hierarchical data model uses a tree (“inverted tree”) structure with hierarchies (Figure 4.2 above). The nodes are objects that are of interest to the enterprise and the branches represent the relationships between the objects. The root node (A) is also referred to, as the parent and child nodes (e.g. B, C and D) are children of the immediate parent (A). The tree is made up of many subtrees such as (A,B,C,D), (B,E,F,I,J) and (D,G,K,H). The model supports only a one-to-many (one-to-one being a special case) relationship between the parent and the children. A parent, therefore, can have many children but a child must have only one parent. The relationships are implemented by using pointers.

#### **4.2.1.2 Limitations of the hierarchical data model**

The hierarchical data model presents some data insertion problems because such actions require the restructuring of the source code of the programs managing the database. The model offers only program-oriented navigational interfaces to the data (Rainer, 2002), making ad hoc queries (i.e. linking data in a way that is not predefined) difficult. The lack of the link between the child nodes in the model results in redundancy of data. For example, if **K** is a supplier to both **B** and **D** (Figure 4.2 above), then the supplier details would be duplicated in the database. Today, this data model is only in use in some legacy systems (Harrington, 2000).

#### **4.2.2 Network data model**

The network data model was an improvement on the hierarchical model that allowed the linking of the child nodes (members) among themselves and their parents (owners), thus, allowing the representation of many-to-many relationships. This enhanced searching capabilities and reduced the amount of data redundancy in databases.

##### **4.2.2.1 Limitations of the network data model**

The network model was just as navigational as its predecessor and still does not fully support ad hoc querying of databases. Complex programs are required to solve simple queries because of the need to trace the linkages among the different data types. Any changes in the data structure necessitated restructuring of the application programs (i.e. there was minimum data independence between the user and the physical storage). Although the network data model led to minimisation of data redundancies, it was dogged by data insertion problems.

#### **4.2.3 Relational data model**

Relational data model was developed to overcome the shortcomings of the network and hierarchical data models such as data redundancy and lack of support for ad hoc queries. The relational data model became widely used for applications that require extensive online transaction processing (OLTP) and ad hoc queries (Harrington, 2000) such as

Geographic Information Systems. Online transaction processing applications are those that are required to input and update large amounts of data into a database continuously.

The relational data structure consists of uniquely identifiable relations, which are physically represented as tables. Relations represent things or objects (entity types) that are of interest to the enterprise and have independent existence such as the fish species, fishing sources and the fishing gear in a fishing organisation. Each relation has a name and is made up of named attributes (columns). Attributes are characteristics associated with entities. The values of the attributes constitutes the rows (tuples) that define the elements of a relation. A fisherman's attributes could be his/her first name, surname, gender and identity number.

### FISHERMAN

FishermanID	FirstName	Surname	Gender	DateOfBirth	MaritalStatus	VillageID
100	charles	Makonese	Male	2/1/69	Married	1
111	Charles	Ruhoto	Male	1/2/72	Single	4
112	Peter	Jamali	Male	12/15/50	Widowed	2
123	Makonese	Tichaona	Male	12/12/72	Single	2
125	Innocent	Hapazani	Male	9/28/65	Single	6
200	Nehemiah	Mavetera	Male	8/15/68	Married	7
305	chipo	Munhu	Female	9/9/78	Single	5
405	Nancy	Chandia	Female	1/18/82	Single	1
505	kudakwashe	paradzayi	Male	4/22/98	Single	1
506	Danai	Paradzayi	Female	2/18/01	Single	4
1000	Zsazsa	Milongo	Female	2/5/75	Married	2

Figure 4.3 An instance of a Fisherman relation

Figure 4.3 (above) shows an instance of a fisherman relation with eleven tuples and seven attributes.

#### 4.2.3.1 Properties of relations

According to Pratt et al, 2000, each distinct relation is characterized by:

- **each attribute having a distinct name**

Each attribute must have a distinct name and must draw values from a particular domain. The ordering of the attributes is irrelevant in a relation. Semantically

sensible comparisons can be carried out (Benyon, 1990) between relations by linking them together through reference to common attributes. For example the VillageID (foreign key) in Figure 4.3 above refers to the VillageID (primary key) in Figure 4.4 below which shows an instance of the Villages relation.

## VILLAGES

VillageID	VillageName	VillageHead	MalePopulation	FemalePopulation
1	Biti Manyanga		830	827
2	Changombe		835	980
3	Galula		1567	1695
4	Gua		1819	1767
5	Ifuko		1822	1895
6	Ifwekenye		2324	2455
7	Isanzu		925	1062
8	Iseche		754	705
9	Iyovyoy		941	962

Figure 4.3 A Villages relation

- **single-valued (atomic) entries**

Each cell of a relation must contain exactly one atomic (single) attribute value, therefore, there can only be a single value at a tuple/attribute intersection (Benyon, 1990). Values of attributes that have multiple values with respect to the primary key must be entered in their distinct rows (see 4.5.2 for an example). This property ensures that the relation is normalized i.e. in first normal form.

- **distinct tuples**

In order to avoid redundant data, no two tuples may contain exactly the same values for **all** attributes, thus, an attribute or a combination of attributes (candidate keys) can be used to identify each tuple. A primary key (single or composite attribute) is the least number of attributes needed to distinguish one tuple from another. The way tuples are ordered in a relation is irrelevant.

### 4.2.3.2 Limitations of the relational data model

The relational data model has limitations in terms of its ability to model real world complex data types and complex relationships e.g. a “parts explosion” (Connolly *et al*,

2002) where an entity is made up of parts, which in turn are also made up of their own individual parts. The overall design of a database application may be complex (due to many relations that may be needed), resulting in slower search and access times when compared to the hierarchical and network models. For the relational model, data integrity has to be enforced by good design principles since it is not inherently part of this model (Rainer, 2002).

#### **4.2.4 Object-oriented data model**

The object-oriented data model seeks to overcome the limitations of the preceding data models when modelling complex real world relationships. This model is useful for applications in heterogeneous environments and those applications that support temporal and spatial dimensions (versioning).

The object-oriented data structure is based on objects and their associated classes. An object is any thing (real or conceptual) that is of interest to the enterprise. Objects are composed of attribute values that represent their state and methods that define their behaviour. Objects are related to their parent or child objects via unique object identifiers (wp\_wdbfits.htm, 2002) similar to the network model in Section 4.2.2. Every object is an instance of a particular class that groups objects which share similar characteristics. A class possesses the qualities of persistence, polymorphism and inheritance so that no matter where it is used, it behaves in a predictable, known way (Harrington, 2000).

- **Persistence**

Persistence is the ability of an object to preserve data across successive program executions, and to even allow such data to be shared by many programs.

- **Polymorphism**

Polymorphism means “many shapes” (Bradley et al, 2001). This means that the same object may be implemented in different environments. For example, the *Print* object can be used to send output to a printer, form or a picture box in Visual Basic.

- **Inheritance**

Inheritance is the ability to create a new class from an existing class. The objects in the new class will have all the characteristics of the objects of the existing class but with additional characteristics to meet the intended needs of the new class. For example, a Person class may be characterised by person *name* and *identity number*. A Student class (with *course* and *department* attributes) may inherit *name* and *identity number* attributes from the Person class to give information about a person who is a student. The Person class is the superclass of the Student class and the Student Class is a subclass of the Person class (Silberschatz et al, 1997).

#### **4.2.4.1 Limitations of the object-oriented data model**

Technological advances make it possible for object-oriented data models to support ad hoc queries. However, Harrington, 2000, argues that the object-oriented model is not well suited to ad hoc queries because it has navigational access to the data. This means that queries must follow predefined paths and inserting new relationships on the fly is not easy.

#### **4.3 Entity-Relationship (ER) Modelling**

Entity-relationship modelling is a top-down approach data modelling technique, which is characterized by the selection of entity types and their associated relationships, and the assignment of their attributes respectively (Howe, 1989). It is based on rules (enterprise rules) representing real world conditions or practices of the enterprise. These rules govern the association between the data elements to be represented in a conceptual and logical data models. These rules enable system designers to clearly identify the channels of information flow within the enterprise.

The conceptual model is depicted using an entity-relationship diagram (ERD), which is a graphical technique for representing data models. Chen's notation, Crow's foot notation and UML (Unified Modelling Language) notations are widely used when drawing up entity-relationship diagrams. A comparative schedule of these three notations is shown in

Figure 4.4 (overleaf). In all notations, entity types are represented by rectangles with the name of the entity inside them. Attributes are depicted using ellipses in Chen's notation, and listed in an attribute rectangle in Crow's foot and UML notations. Primary keys are underlined using solid lines and foreign keys underlined using dotted lines in both Chen's and Crow's foot notations. In UML notation, {PK} and {FK} may be used to indicate primary attributes and foreign keys respectively.

### 4.3.1 Relationships

Relationships represent associations or logical links between two or more entity types (Benyon, 1990), for example, between fisherman and fishing gear. Chen's notation uses a diamond shape to depict relationships. A line with the name of the relationship is used in both Crow's foot and UML notations. UML notations also indicates the direction of the relationship.

The degree of a relationship is the number of entity types (participants) that participate in that relationship. A unary (recursive) relationship occurs when the same entity type participates in a relationship more than once in different roles e.g. one fish species consumes other fish species. A binary relationship is one that involves two entity types e.g. the association between a fish species and a particular river. Three or more entity types participating in a relationship at once results in ternary or n-ary (n representing the actual number of participants) relationship respectively.

### 4.3.2 Cardinality of a relationship

Cardinality describes the maximum number of possible relationship occurrences for an entity participating in a relationship resulting in one-to-one, one-to-many or many-to-many relationships between entity types (See Figure 4.4 overleaf). In a one-to-one relationship, a single entity occurrence from one entity type is related to a single entity occurrence from the other participating entity type. For example, a fisherman may have only one fishing license. In a one-to-many relationship, a single entity occurrence from one entity type is related to many occurrences from the other participating entity type. A typical example is the case of a village hosting many fisherman, but each fisherman

staying only in a particular village. A many-to-many relationship involves multiple associations between the entity occurrences for both participating entity types. A fisherman may catch many types of fish species, but a particular fish species may be caught by many fisherman.

Construct	Chen	Crow's foot	UML
Entity			
Attribute			
Relationship			
Recursive relationship			
<b>Cardinality</b>			<b>Has</b>
One-to-One 1:1			
One-to-Many 1:M			
Many-to-Many M:N			
<b>Participation</b>			
Mandatory on both sides e.g. 1:M			
Mandatory on one side			
Optional on both sides			
<b>Generalisation or Specialisation</b>			

Figure 4.4 Entity-relationship diagram notations

### **4.3.3 Participation constraints**

Participation constraints determine whether all (mandatory participation) or only some (optional participation) entity occurrences participate in a relationship. This results in either mandatory participation on both sides, mandatory or optional participation on one side, or optional participation on both sides of the relationship (Figure 4.4 above). For example, if given that a fisherman must stay in a village and that a village must have at least one fisherman, then the resulting relationship is mandatory participation on both sides. A relationship that arises when a fisherman may catch many kinds of fish species but a particular fish species may be caught by many fisherman, has optional participation on both sides.

### **4.3.4 Generalisation/Specialisation**

Specialisation highlights differences between entity types by defining one or more subclasses of a superclass entity type. Student and Lecturer entity types (4.2.4) are specialisations of the Person entity type in a university organisation. A student inherits all the attributes of the person entity but has other attributes that pertain that Student entity type only and not to the Lecturer entity type. Generalisation is the reverse of specialisation. Generalisation identifies common features between entity types to define a generalising superclass entity.

## **4.4 Connection Traps**

During the data modelling process, connection traps occur when three entity types are related to each other. Fan traps and chasm traps result in indeterminate associations or the non-existence of the path respectively, for certain entity occurrences between two of entity types participating in the relationship (Benyon, 1990). Connection traps can be overcome by introducing a relationship between the disconnected entities.

### **4.4.1 Fan traps**

Fan traps occur when a many-to-one (M:1) relationship is immediately followed by a one-to-many (1:M) relationships such as the one shown in Figure 4.5 (overleaf). The

problem in this example is to uniquely associate the employee with his/her department when there are many departments in a division, and each division has many employees.

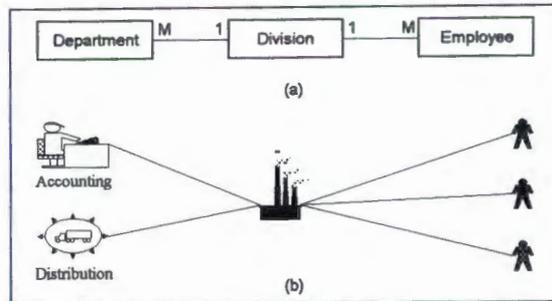


Figure 4.5 Fan trap

The fan trap is resolved by establishing a one-to-many relationship between the disconnected entity types (such as Employee and Department) as indicated in Figure 4.6 below.

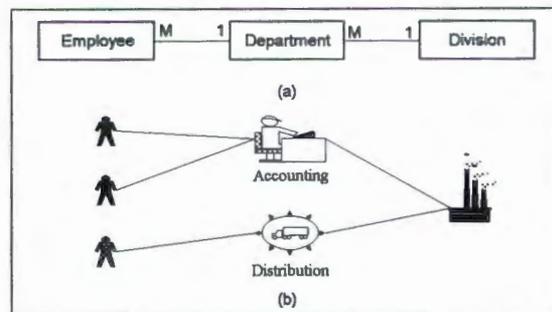


Figure 4.6 Resolving fan traps

#### 4.4.2 Chasm traps

Chasm traps can occur when there is an optional participation of entity types involved in relationships. As a result, the path may not exist across the entities for certain entity occurrences. Figure 4.7 represents the facts that a single river must serve at least one or more fisherman who may catch zero or more fish species. A particular fish species may be caught by zero or one fisherman. The problem in this case is that it is not possible to determine which fish species is found in which river.

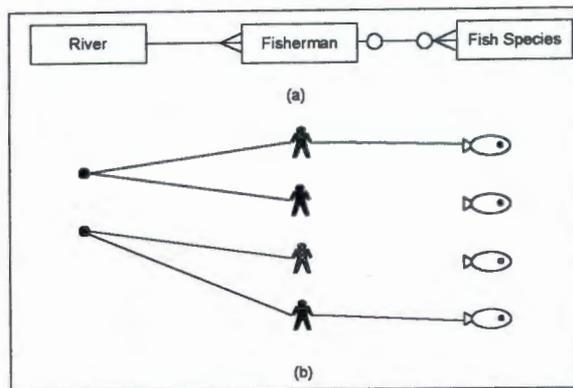


Figure 4.7 Chasm trap

This is resolved by establishing a one-to-many relationship between the river and fish species entity types as shown in Figure 4.8 below.

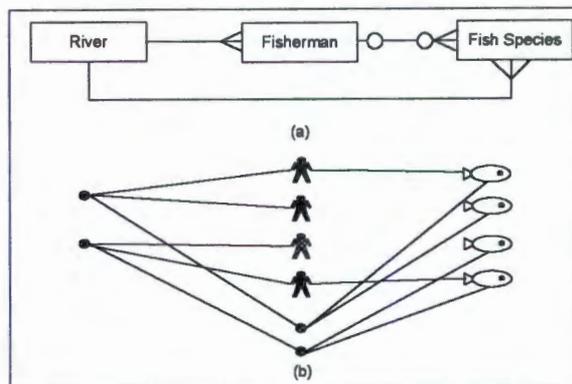


Figure 4.8 Resolving chasm traps

## 4.5 Normalization

Normalization is a formal process used by database designers to identify the optimal grouping of attributes for relations based on their primary or candidate keys and the functional dependencies among their attributes to prevent possible occurrence of anomalies such as update (modification), deletion and insertion (Connolly et al, 2002).

### 4.5.1 Functional dependency

Functional dependency is about the meaning of data within the context of the enterprise. It describes the functional relationships among the attributes of a relation (Atzeni *et al*,

1999). Attribute B is functionally dependent on attribute A (denoted  $A \rightarrow B$ ), if each value of A is associated with exactly one value of B (4.5.3). Attribute A is referred to as the determinant of B. Transitive dependency arises when three or more attributes are related to each other (e.g. A, B, C) such that  $A \rightarrow B$  and  $B \rightarrow C$ . As a result, attribute A determines attribute C through attribute B (4.5.4).

#### 4.5.2 First normal form (1NF)

The 1NF allows dealing only with atomic values in a relation (4.2.3.1) i.e. each attribute has only one value. It disallows multivalued attributes, composite attributes and their combinations to be stored in a relation.

Name	Fish species	Gear	CatchSite	Date	Weight	Site Code	Source
Charles	Barbus	MB11	Makonese	01/02/02	50	LA1	Lake Rukwa
	Clariidae	NT10	Marufu	01/03/02	15	R1	Save
Peter	Barbus	NT10	Mudonhi	10/05/02	25	R1	Save
	Barbus, Clariidae	CN01	Musekiwa	10/06/02	12,36	R2	Chinyika

**(a) Unnormalised relation**

Name	Fish species	Gear	CatchSite	Date	Weight	Site Code	Source
Charles	Barbus	MB11	Makonese	01/02/02	50	LA1	Lake Rukwa
Charles	Clariidae	NT10	Marufu	01/03/02	15	R1	Save
Peter	Barbus	NT10	Mudonhi	10/05/02	25	R1	Save
Peter	Barbus	CN01	Musekiwa	10/06/02	12	R2	Chinyika
Peter	Clariidae	CN01	Musekiwa	10/06/02	36	R2	Chinyika

**(b) Relation in First normal form (1NF)**

Figure 4.9 Decomposing to First Normal Form

#### 4.5.3 Second normal form (2NF)

The 2NF is based on the concept of full functional dependency, where every attribute is functionally dependent on the primary key (Benyon, 1990) only. When the primary key consists of several attributes, no subset of the primary key should determine the value of a non-key attribute. All non-key attributes shall be dependent on all the primary keys (i.e.

the value of a non-key attribute cannot be learned from knowing values for only part of the primary keys). Figure 4.10 (a) shows a River Catch entity type with the attributes, Fish species, Weight, SourceID, Source name and Length. The primary key consists of the Fish species name and SourceID. The Source name and Length is determined only by the SourceID but the Weight is determined by both the Fish species and SourceID. To conform to 2NF, Source name and Length should be moved to a separate entity with SourceID as the primary key as shown in Figure 4.10 (b).

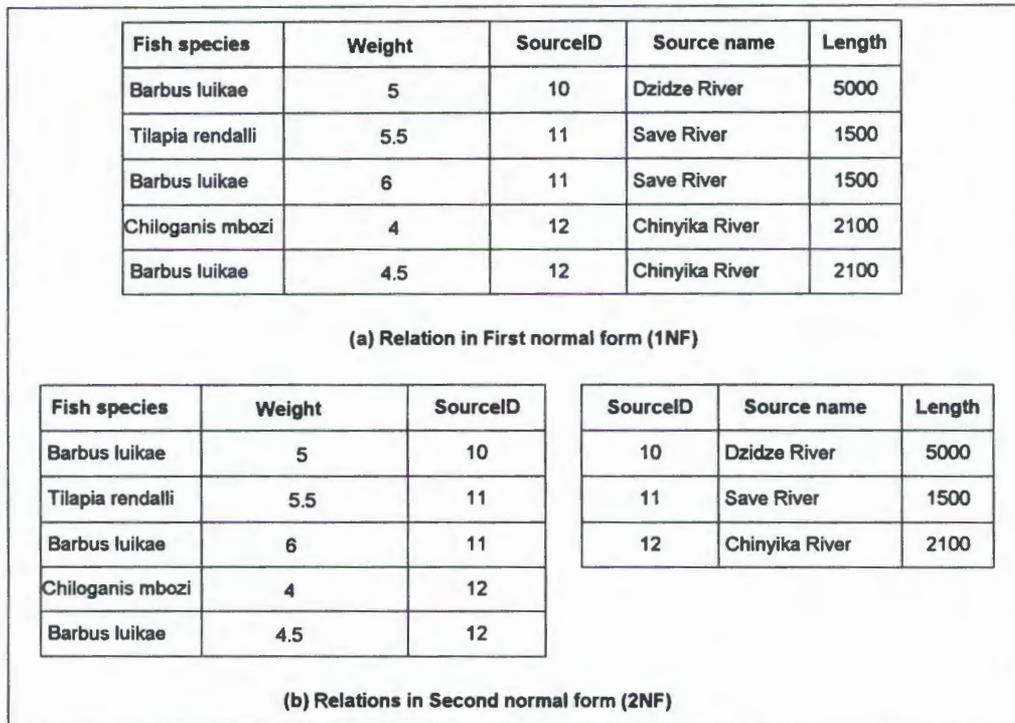


Figure 4.10 Decomposing a relation to 2NF

#### 4.5.4 Third normal form (3NF)

3NF is based on the concept of a transitive dependency, where a non-key attribute is dependent on another non-key attribute. For 3NF, no nonkey attribute should determine the value of another non-key attribute. The value of any non-key attribute should depend only on the primary key. The existence of transitive dependencies within a relation gives rise to redundancy in a data model, with potential update anomalies.

Figure 4.11 below shows how the Fisherman relation with NID as the primary key. The problem here is that attribute VillageName is dependent on VillageCode attribute which is a non-key. The VillageCode is dependent on the NID attribute, therefore, transitive dependency exists because NID can determine VillageName through VillageCode. To conform to 3NF, a separate relation should be defined to correlate the VillageCode and the VillageName.

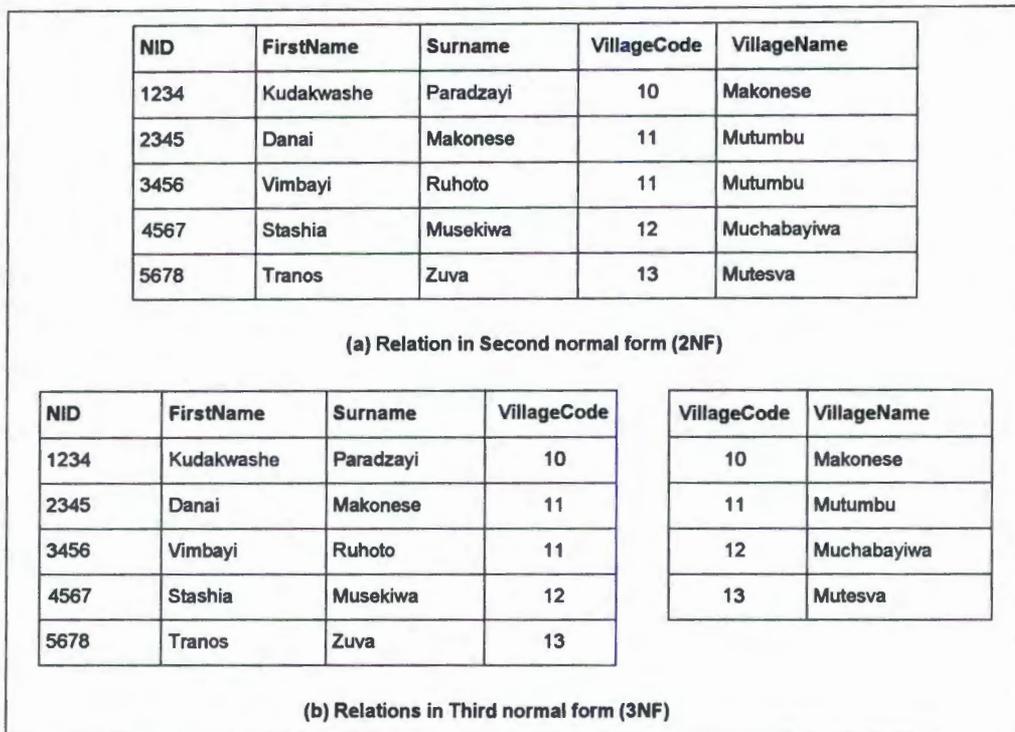


Figure 4.11 Decomposing to 3NF

#### 4.5.5 Boyce Codd Normal Form (BCNF)

Boyce Codd Normal Form is based on functional dependencies that take into account all the candidate keys in a relation. Relations are in BCNF if in each table, every determinant (4.5.1) is a candidate key (Benyon, 1990). A candidate key is any key that can uniquely identify a tuple in a relation and, if selected, it becomes the primary key of the relation. BCNF deals with candidate keys and not just primary keys. If a relation is in BCNF then it satisfies the requirements of the other (First, Second and Third) normal forms.

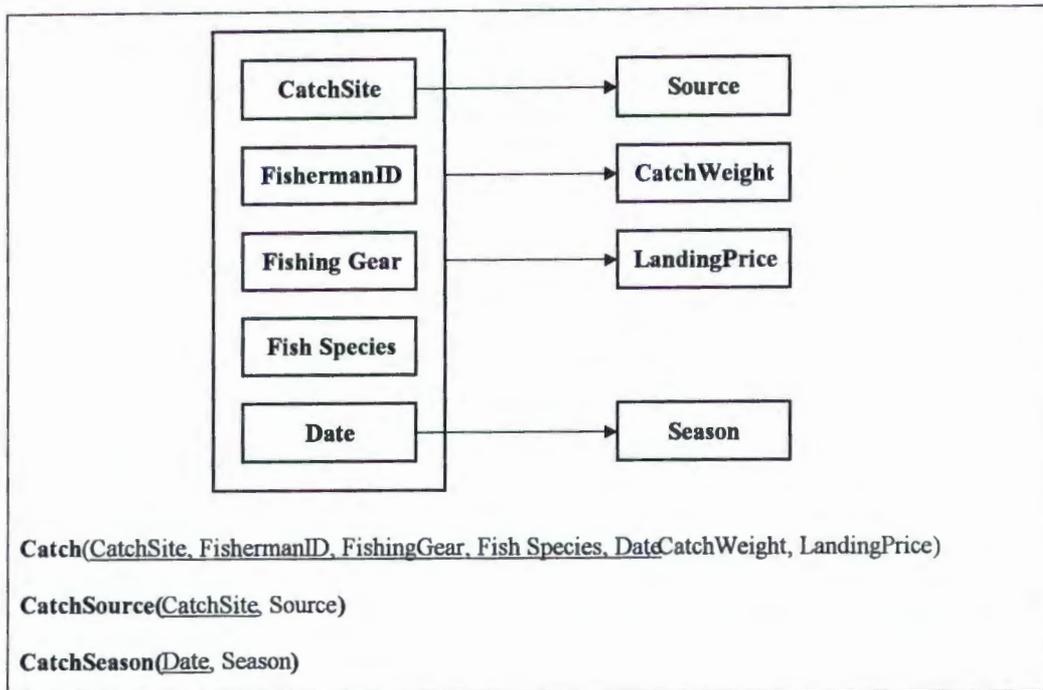


Figure 4.12 Functional dependency diagram

The BCNF is derived directly and automatically from functional dependency diagrams as illustrated in Figure 4.12 above. The diagram represents attributes that constitutes a River catch in a fishing organisation. The catch site is determinant of the fishing source, and the season is functionally dependent on the catch date. (Catch site, fishermanID, fishing gear, fish species and date) values directly determine the catch details such as catch weight and landing price for a particular fish species. BCNF requires that each determinant be a candidate key in a relation. Therefore, three relations are required to represent the data.

#### 4.6 Conclusion

This chapter gave a terse account of the fundamental concepts for sound data modelling techniques. Data modelling minimises redundancy resulting in consistent data in a database application.

## 5 LRBIP-EIS USER NEEDS ASSESSMENT

The view integration approach (Section 3.3) will be used to perform the user needs assessment of the Lake Rukwa Basin Integrated Project Environmental Information System (LRBIP-EIS) database. This approach identifies the information requirements of each sectoral user (e.g. fisheries, apiculture, wildlife) separately. The separate data models are then integrated into a global data model that will be used for designing the database application.

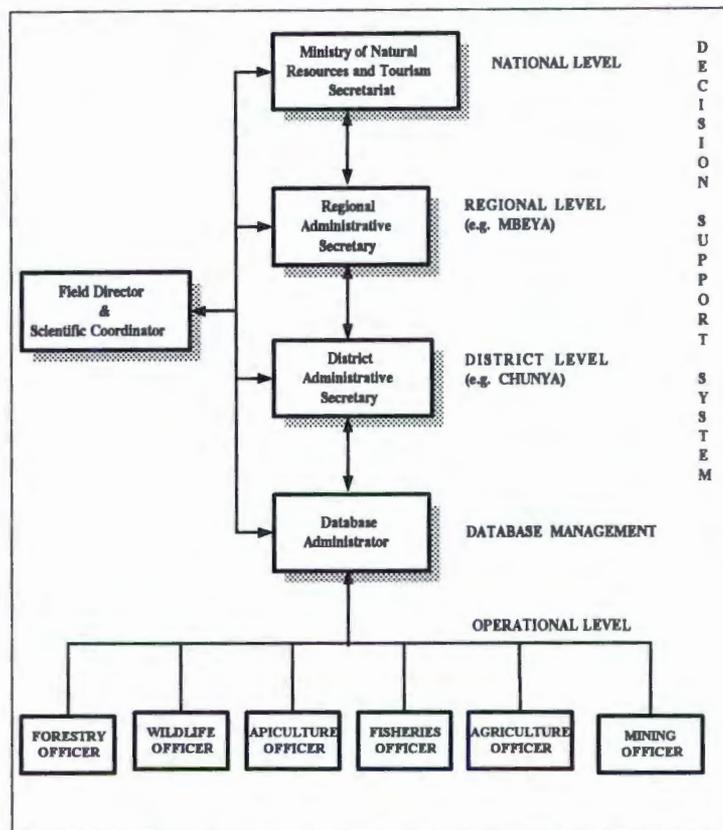


Figure 5.1 Lake Rukwa Basin Integrated Project Organisational Chart

The user needs assessment of the Lake Rukwa Basin Integrated Project environmental information system has to be carried out within an institutional framework. The environmental information system should collect information toward the fulfilling of the objectives of environmental management plans at both decision-making and operational

levels. It should therefore, provide for the information needs of the secretariat (at ministerial, regional and district level), environmental scientists/managers in both public and private sectors, environmental management councils (e.g. fisheries, wildlife) and members of the general public with interest in environmental matters. Figure 5.1 above, shows an adaptation of the national framework outlined in Section 1.2. This framework was proposed after reviewing the literature available on the case study area. It provides for the flow of information between field officers and policy-makers in the Ministry of Natural Resources and Tourism.

- **National Level**

The national level represents the ministry responsible natural resources and tourism. The ministry formulates policy and legislation on the sustainable use of natural resources in the country and also coordinates national environmental information system (EIS) initiatives. The ministry links with donors who fund EIS programs such as the Italian government in the case of Lake Rukwa Basin.

- **Regional Level**

The regional level coordinates district EIS initiatives and report to the Secretariat in the Ministry of Natural Resources and Tourism.

- **District Level**

The district level consists of the administration and operational sections. The administration involves coordinating the effort of various departments with regard to environmental information gathering. The operational sector involves departmental officers working with the local communities to implement environmental policies. The officers give expert advice as well as gathering and processing raw data that will be used to populate the LRBIP-EIS database.

- **Database Administration and Scientific Coordination**

The Database Administrator will be responsible for maintaining the LRBIP-EIS database. He checks that the input data from the various sectors meet a certain

minimum criteria so that it can be shared in the integrated system. The Field Director and Scientific Coordinator are technical consultants who liaise with all the stakeholders who are involved in the design and implementation of LRBIP-EIS, from Ministry officials right down to the field officers responsible for different environmental sectors.

### **5.1 Components of an Environmental Information System**

The integrating nature of the Lake Rukwa Basin Integrated Project requires development of a harmonised information system that caters for the requirements of each individual environmental sector shown in Figure 5.1 above. The Food and Agriculture Organization of the United Nations (FAO) has identified six components (natural, biological, technological, economic, socio-cultural, institutional and legal) that make up information systems for managing and monitoring the sustainable utilization of natural resources (Crispoldi, 2001).

- **Natural component**

This component deals with environmental and physical factors that affect the fauna and flora species population, reproduction and migration behaviour. The factors include seasonal and climatic conditions, rainfall and temperatures, and species composition and density.

- **Biological component**

The biological component deals with the biological characteristics of all the fauna and flora species found within a given management area such as growth rate, size, age, mortality, inter species relationships.

- **Technological component**

The technological component deals with the technological equipment that is used to exploit natural resources in each environmental sector. For example, in the fisheries sector, the equipment includes vessels and fishing gears. There is need for a systematic way of recording the hardware inventory to make it possible to

identify and track fisherman and fishing vessels (and gear) performance through time regardless of changes in vessel ownership, location or fishing activity (NOAA, 2002).

- **Socio-economic component**

The socio-economic component covers the social and economical aspects of an environmental sector. It should provide information on the number and social status of the people involved in that sector, including occupational mobility; and individual income and income distribution; production levels; harvesting and post-harvesting costs; and market prices.

- **Institutional and legal component**

This component deals with the institutional and legal issues of each sector, clearly defining the institutional frameworks for data collection and dissemination. It deals with the licensing requirements to enable individuals to operate legally. The central government can devolve licensing powers to local management committees to empower communities to develop a sense of ownership of resources and thus, encourages them to take greater responsibility in the conservation and protection of the natural resources.

The information gathering guide (**Appendix A**) will be used to carry out the user needs assessment for six sectors (fisheries, apiculture, wildlife, agriculture, forestry and mining) of the Lake Rukwa Basin Integrated Project. The assessment will identify spatial and non-spatial information requirements for each sector with regards to the components of an environmental information system outlined above.

## **5.2 Fisheries**

In Tanzania, the Fisheries Act of 1970 is the principal legal instrument for the administration of the fisheries sector. The Act provides for the protection, conservation, development, regulation and control of fish, fish products and aquatic flora (Ssentongo *et al*, 2000). Among the supporting pieces of legislation, amendments government notice

(G.N) 370 of 1994 and G.N 189 of 1997 are of particular relevance as they prohibit the use of certain specified vessels or tools and beach seines in fresh water fisheries. In most industrial and semi-industrial fisheries, the submission of data on fish harvested is a compulsory requisite of the license. The information is supplied through logbooks designed to collect data on the time and place of fishing, fishing effort and catch by species. Information may also be supplied through catch sampling programs at the fishing source or landing site. However, the collection of data is fragmented in artisanal (small-scale) fisheries, which consist of large numbers of small and dispersed units operating on subsistence or semi-commercial level. In Tanzania, this problem has been worsened by the ongoing retrenchment and restructuring process within the public sector resulting in limited coordination for data collection, entry and submission at the regional level. Most of the field enumerators have been redeployed on other duties and are no longer exclusively involved with data collection for fisheries sampling programs (RFIS, 2001).

Although, fishing is the main activity in the Lake Rukwa Basin, Bianchini, 2001, reckons that the knowledge on the Rukwa ichthyofauna is still limited. Fish stocks are under considerable threats from overexploitation, including indiscriminate fishing practices such as the use of gill nets and poison to kill fish, habitat destruction through pollution and siltation of watercourses and lakes, lack of sound policy on the issuing of fishing permits, and shortage of manpower to enforce fishing regulations (Mukwada, 2000). These problems have resulted in catches reduced in quantity and fish size (Rüther, 2001). According to Davenport, 2000, there is no official inventory of fish species, distribution, population and harvesting quotas. Fish production information such as annual fish yields, yield per fish species, and the number of fisherman and canoes, is highly inconsistent in the case study area.

There are fishing gathering centres at Maleza and Somang'ombe, where the fishermen dispose of their catch to fish traders. Fish traders then distribute the fish in markets located in places such as Mbeya Municipal, Mbeya Rural, Rungwe, Kyela, Mbozi, Tunduma and Chunya. Only a minimal fraction of the catch is sold fresh, the rest is sun-dried, salted or hot-smoked before it is dispatched to the markets (Bianchini, 2001). A

large proportion of the fish is reportedly smuggled to foreign markets such as Zambia and Congo (Rüther, 2001). Schools, institutions, villages and private individuals that are distant from the lake undertake fish farming (aquaculture) for subsistence or limited commercial gains on small and medium basins (CIC Report, undated).

### **5.2.1 Fisheries information requirements**

The following information requirements are drawn from the analysis of the fisheries components above. Crispoldi (2001) recommends a minimum set of data internationally regarded as essential for national fisheries management and international comparability purposes.

Fisheries spatial data requirements include:

- Fish species distribution on lakes and rivers (breeding and fishing grounds)
- Lake and river networks
- Fisherman village distribution

The fisheries component should provide attribute data on the following:

- Biological characteristics of fish species
- Personal details of the fishermen and fish traders
- Estimates of total fish production, by species, gear and area of capture, fisherman. (Combined with data on size at capture, permit analyses relating to gear selectivity and indices of exploitation).
- Estimates of fisheries income [total value of fish production by species; income per fisherman; income per village]
- Average producers price at landing [Generate time series and trends for use in economic analysis and market studies]
- Fishing effort by fishing grounds and species constitute elements in the formulation of indices of abundance [time series of fishing effort are indicative of declining or increasing trends of fisheries performance in districts and regions]
- Daily and monthly (raised estimate) summaries for each landing site by fishing category (e.g. hook and line, net, long lines, traps, beach seine).

- Monthly district (raised) summary for each fishing category (by species, weight and value) and monthly summary across the district by fishing category
- The number and type of fishing vessels and gear, and their size and capacity, is a measure of the level of effort put into the fishery [combined with catches and hours fished provides an indication of the performance; comparison of licensing requirements against actual catch gives an indication of compliance to restrictions]

### 5.3 Apiculture

Apiculture is the keeping of bees on a commercial scale for the production of honey. The abundance of the ‘miombo’ type of vegetation (Picture 5.1 below) in the Lake Rukwa region creates favourable conditions for beekeeping. Although annual honey production and revenue figures are not available, it is estimated that 20 tons of honey and 5 tons of wax were harvested per annum over the last few years (CIC Report, undated). There are two types of bee species that are found in the Lake Rukwa Basin, the *mellifera scutelata* (very aggressive) in the plains and the *mellifera monticola* (less aggressive) in mountainous regions.



Picture 5.1 ‘Miombo’ forest

(Source: <http://www.rupertshoney.co.za/rh/beekeeping%20images.html>)

Lack of licensing procedures makes it difficult to keep track of the beekeepers that are permitted to trap bees and collect honey in Forest Reserves such as Muipa and Lukwati. The beekeepers rely on traditional methods for the collection of honey, and in most cases use fire to inactivate the bees. This act kills the bees subsequently affecting the size and number of swarms. On occasion, these fires run out of control and spoil the environment (Hassan, 2001).

Cylindrical beehives are constructed from tree barks obtained through a process of ring barking larger trees. Hassan (2001) points out that up to five trees are destroyed in the process of making a single bark hive, which has an average lifespan of 2.5 years. The shortage of large trees suitable for bark hive making is already evident in the region. The unselective tree felling for tobacco farming, fuel wood demands, and tsetse fly control for livestock is compounding this problem further (Hassan, 2001).

### **5.3.1 Beekeeping information requirements**

The spatial data requirements for the apiculture sector include :

- Location of apiaries (beehive sites) and forest camps
- Road network
- Market locations
- Land use/land cover data including forest cover
- Extent of fires
- Distribution of diseases affecting apiaries

The beekeeping component should provide attribute data on the following:

- Beekeepers and honey traders
- Honey and honey-related product harvests
- Climatic conditions (temperature, rainfall, blooming seasons)

The users require the system to provide for analyses such as:

- Compare wildlife poaching activities relative to apiary site locations

- Compare honey production relative to surrounding land use/land cover including forested areas
- Provide time series of harvest and income generated for each beekeeper, village and district.

#### **5.4 Wildlife**

The case study area is endowed with a diverse range of animal fauna. The prominent mammal fauna include Kudu Major, the Lichtenstein elk, the equine antelope, the black antelope (miombo woodlands), the topi, the cane field antelope (grasslands around the lake), elephant (water basin), hippopotami, and crocodiles (lake and neighbouring swamps). The bird fauna in the case study area is also very diverse with at least 360 species, including among the most important ones, the robust parrot (CIC Report, undated).

Wildlife activities are governed by the Wildlife Conservation Act of 1974, which is administered by the Ministry of Natural Resources and Tourism. This piece of legislation by default, prevents interaction between the indigenous rural populations and the protected areas adjacent to the rural territories. Local people may not hunt wildlife unless they obtain a hunting permit and are then allocated their quota at a fee. According to Musendo (2001), the fee can be as high as USD 7500.00 per annum. As a result, subsistence hunting is illegal and poaching of wild animals by the local population is rife. Legal hunting is carried out by Safari operators (catering mainly for the international clientele) and local hunters (predominantly expatriate residents and wealthy nationals) who can afford to purchase hunting licenses from central government (Davenport, 2000).

There is no current inventory of hunted mammals and birds (hunting quotas, licensing, zones, seasons, species [herbivores or carnivores], seasonal routes etc). There is only limited monitoring of animal killings in the game reserves and game controlled areas. As a result, there is insufficient quantitative data to use for the determination of annual hunting quotas. Given the importance of the income in foreign currency, which derives from hunting concessions, a scientific system should be introduced for the definition of

sustainable quotas over the long term, for each of the protected areas (CIC Report, undated). From the same Report, it is believed that the wild species (some of them endangered e.g. puku) in the region of Mbeya are diminishing drastically. This is highly speculative because there is no statistical evidence available on the processes of loss of any single species.

There is limited quantitative data on the impact of animals described as 'vermin' in some local reports. These animals (e.g. baboons, elephants and bush pigs) inflict crop damage (Mbassa *et al*, 2001) while hippopotami, crocodiles and lions cause loss of both human and domestic animal lives. The government authorities usually kill these troublesome animals for trophies through a culling process.

Wildlife activities are closely related to the tourism sector. CIC Report (undated) suggests the possibility of cultural villages (host masters and/or tour leaders) with scheduled hunting and photography visits that are centrally coordinated. This would be ideal if animal seasonal routes can be identified and visits scheduled accordingly. However, the lack of quantitative data on the large mammal populations and their seasonal movements through the entire eco-system creates a serious limitation in the design and implementation of appropriate plans of action.

#### **5.4.1 Wildlife information requirements**

The information requirements of the Wildlife component are divided into spatial data and attribute data components.

##### **Spatial data**

- Species distribution including seasonal migration routes
- Eco-tourism locations (game reserves and hunting grounds, animal viewing sites)
- Spatial distribution of poaching activities
- Administrative boundaries
- Infrastructure (roads, transmission lines, telephone lines, health facilities)

### **Attribute data**

- Species composition and abundance
- Hunting activities including licensing, quota determination and allocation, monitoring hunting harvests and overtakes, checking compliance by reconciling hunting licenses and animals harvested (Debriefing Note, 1999)
- Hunting clubs and Safari tour operators
- Poaching activities giving vulnerable and threatened species, details of confiscated weapons and records on any criminal convictions
- Record of government trophies with records of origin/source (whether it is from vermin or confiscated from poachers)
- Quantitative data on vermin [specific areas trends with regards to human and livestock losses, damage to crops (Davenport, 2000)]

### **5.5 Agriculture**

The main ethnic groups residing in the Lake Rukwa Basin are the Wabungu, Wanyamwezi/Wakonongo, Wanyamwanga (sedentary groups) and the Wasukuma (Mbassa *et al*, 2001). The Wasukuma are semi-nomadic, and practice semi-sedentary agriculture (“shifting cultivation”) with a rotation cycle in the zones under cultivation of about three years or so. Shifting cultivation is associated with the slashing and burning of large amounts of vegetation. The sedentary groups engage in subsistence farming for food crops, and limited commercial production of tobacco and cotton. However, tobacco farming makes use of a drying and curing process, which relies on the use of large amounts of wood, impacting negatively on the population of large trees (Rüther, 2001). The settled farmers keep large quantities of cattle, goats and sheep further straining the available vegetation resources.

The Wasukuma are involved in pastoralism. They use thin bands of territory that are not infected by the tsetse fly, to reach the watering points along Lake Rukwa and new pastures in the area bordering Zambia. During their seasonal treks, they inevitably compete with settled farmers for grazing pastures resulting in overgrazing and deforestation (CIC Report, undated).

### **5.5.1 Agriculture information requirements**

The spatial data requirements of the agriculture sector include:

- Land use/land cover giving indications for agricultural zones and grazing areas, including areas of conflict
- Spatial distribution of tsetse fly infested areas
- Distribution of veld fire incidences
- Locust breeding areas (grasslands with poor drainage)
- Distribution of vermin
- Soil type data

The following attribute data is required for the agriculture sector.

- Crop production and marketing statistics for cash crops (tobacco and cotton) and subsistence crops (e.g. maize, groundnuts, cassava, sweet potatoes)
- Farmer details for those involved in sedentary agriculture
- Climatic conditions (temperature, rainfall)
- Livestock details and statistics
- Quantitative data on vermin [specific areas trends with regards to human and livestock losses, damage to crops (Davenport, 2000)]

### **5.6 Forestry**

Forestry management covers forestry and vegetation related activities. The vegetation found in the case study area is classified as the miombo woodland/savanna type. In the forestry reserves, there appears to be no systematic survey of (forest and timber) quantities, regeneration capacities, growth rate and extractive possibilities of the reserve and the impact of the use of exotic species. There is no inventory of tree species as well as other plant taxonomy. The inventory should have a description and phyto-sociological mapping of flora and vegetation i.e. types, environmental and territorial distribution, and their content in peculiar floristic elements (CIC Report, undated). Information on the exploitation of forest products and their marketing is not readily available at the moment.

Charcoal production is a major activity in the study area. Although the charcoal production license stipulates the tree size, type, quantity and location (Ruther, 2001); enforcement of these restrictions is not thoroughly administered. The high demand of forestry products by the agricultural and apicultural sectors promotes the illegal harvesting of forest products by the local population. Information on the deforestation levels and vegetation characteristics is required in order to gauge the rate of degradation and to ascertain the need for conservation strategies.

#### **5.6.1 Forestry information requirements**

The forestry component should:

- Include an inventory of commercial forest species and distribution
- Provide production and marketing statistics
- Keep track of forestry and charcoal licensing details to monitor compliance
- Provide information on other forestry products e.g. fruit, grass, mushrooms, firewood

#### **5.7 Mining**

Open pit excavation and panning are the main methods practiced by the locals for mining gold deposits. These methods cause ugly surface scars and siltation and pollution of watercourses respectively (Rüther, 2001), posing serious threats to fisheries and human lives. No information on the possible use of cyanide was available. A large percentage of the mining operations are illegal and tree loss is inevitable as the gold seekers clear the land for makeshift homes and wood fuel.

##### **5.7.1 Mining information requirements**

The proposed mining component of the LRBIP-EIS should provide information on:

- Distribution and extent of mining activities by mining methods, minerals, legal and illegal operations
- Production and marketing statistics
- Pollutant levels and watercourses (hence, fish species) under threat
- Extent of forests decimated by mining operations

- Licensing of mining operations

## 5.8 GEODATABASE ACTIVITIES

Generic geographic data underpins the data requirements for all other sectors. The geographic data includes:

- **Boundaries** (villages, game and forest reserves, administrative boundaries (districts, regions, etc))
- **Infrastructure** (Roads [classified], rail, transmission lines, telephone lines, schools, health facilities etc)
- **Topography** (Contours, DTM for slope and landscape analysis)
- **Physical resources** (land use [including census data and livestock], soils, vegetation, geology, geomorphology, climate [including rainfall and evapotranspiration])
- **Hydrographic (water-related) features** (river system, drainage basins, watersheds, lakes and dams, wetlands etc)
- **Conservation and recreation** (areas of critical environmental concern, recreational spaces, scenic landscapes, historic sites which can attract eco-tourism)

## 5.9 CURRENT COMPUTER SYSTEM CONFIGURATION

The current computer hardware and software configuration includes :

- One personal computer (Compaq, 10 Gigabytes, modem, CD-drive, 17 inch screen)
- Telephone (also serves as access to the internet)
- One photocopy machine (Rüther, 2001)
- Recommended software include ArcView and its related extensions (Spatial Analyst and 3D Analyst) for GIS, and Microsoft Access for database development

## 5.10 COOPERATION WITH OTHER AGENCIES/INFORMATION SYSTEMS

The 'Deutsche Gesellschaft für Zusammenarbeit' (GTZ) is engaged in the 'Katavi-Rukwa conservation and development programme' in the Katavi National Park and the

Rukwa Game Reserve. The latter, borders on the Lukwati Game and Forest Reserves, which fall into the LRBIP area. The Tanzanian Parliament declared Lukwati a Game Reserve and joined it to the Rukwa Reserve resulting in a situation where one half of the new Lukwati/Rukwa reserve falls into GTZ project area and the other half is part of LRBIP. This offers an ideal opportunity for co-operation between GTZ and LRBIP in terms of data sharing.

### 5.11 Conclusion

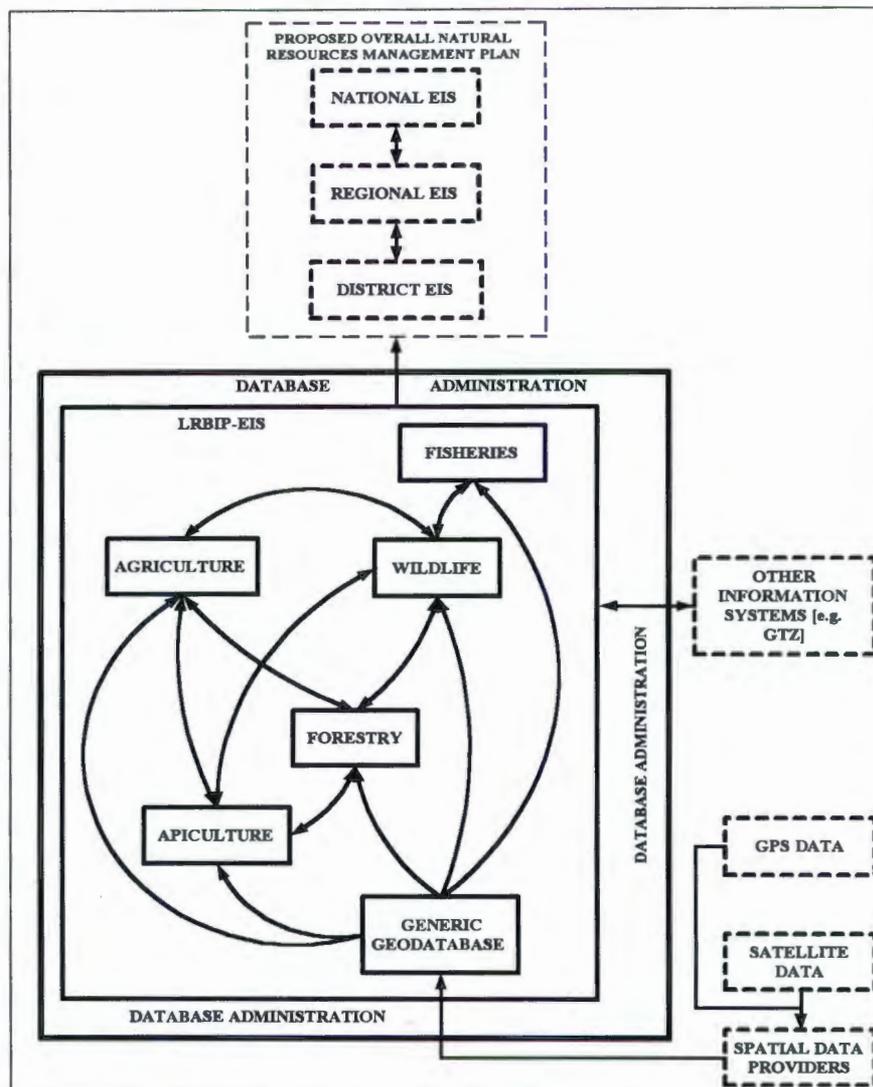


Figure 5.2 Context diagram for the LRBIP-EIS database

Analysis of the user requirements reveals the information flow between different sectors, allowing cross-sector analyses. For example, wildlife poaching activities can be compared against apiary locations to investigate any correlation between them. Figure 5.2 above, shows the information flow between the components of the proposed system.

The system will be able to supply information to central government by uploading data to proposed district, regional and national EIS databases. The generic geographic data will be supplied by spatial data providers such as Infobridge and TANRIC. The cooperation between LRBIP and GTZ should be encouraged in terms of database development to manage the common area linking the two projects.

## 6 LRBIP-EIS SYSTEM DESIGN

The system design phase translates the users' information requirements into the data model upon which the database application will be based. A modular database design approach will be used to develop the separate conceptual and logical data models (fisheries, apiculture and wildlife sectors) for a prototype system. The separate logical data models will then be merged to create the global data model for the LRBIP-EIS database prototype.

### 6.1 Modelling tools

A data model consists of a number of symbols that are linked up according to certain conventions (Bhunu, 1999). The Crow's foot notation (Section 4, Figure 4.4) will be used to annotate entity-relationship diagrams (ERD) in this thesis.

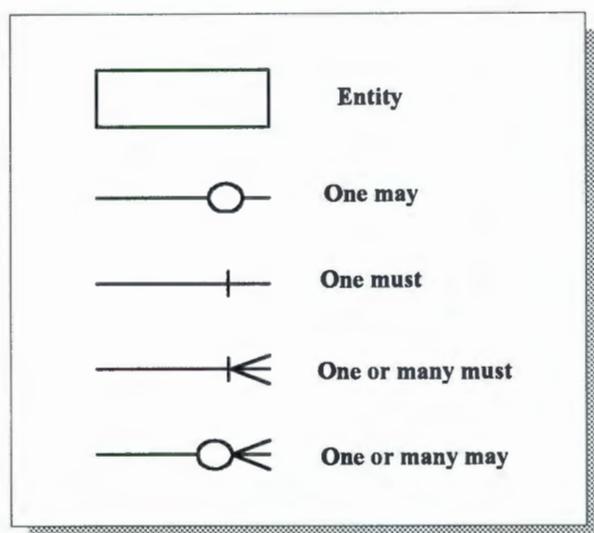


Figure 6.1 Symbol legend

### 6.2 LRBIP-EIS conceptual data modelling

The conceptual data model is constructed from the enterprise rules drawn up from the user requirement specifications. The conceptual model represents the basic entities and their associated relationships using entity-relationship diagrams (ERD). Enterprise rules and an entity-relationship diagram will be drawn up for each sector.

## 6.2.1 Fisheries conceptual data model

### Fisheries Enterprise Rules

#### **Fisherman - Fish Species**

A fisherman may catch many kinds of fish species

A fish species may be caught by many fisherman

#### **Fisherman - Fishing Gear**

A fisherman may own many fishing gears

A fishing gear must be owned by one fisherman

#### **Fisherman - Fish trader**

A fisherman may sell fish to many fish traders

A fish trader may buy fish from more than one fisherman

#### **Fish species - River**

A river may have one or many types of fish

A fish species may be found in many rivers

#### **Fish species - Lake**

A lake may support many types of fish

A fish species may be found in more than one lake

#### **Fisherman - Village**

A fisherman must reside in a village

A village (settlement) may have many fisherman

#### **Fisherman - Wildlife species**

A fisherman may poach many kinds of wildlife species

A wildlife species may be poached by more than one fisherman

A fisherman may be attacked by more than one type of wildlife species

A wildlife species may attack more than one fisherman

#### **Fish species - Wildlife species**

A fish species may be consumed by more than one wildlife species

A wildlife species may eat more than one type of fish species

The enterprise rules are represented in the entity-relationship diagram (ERD) shown in Figure 6.2 overleaf.

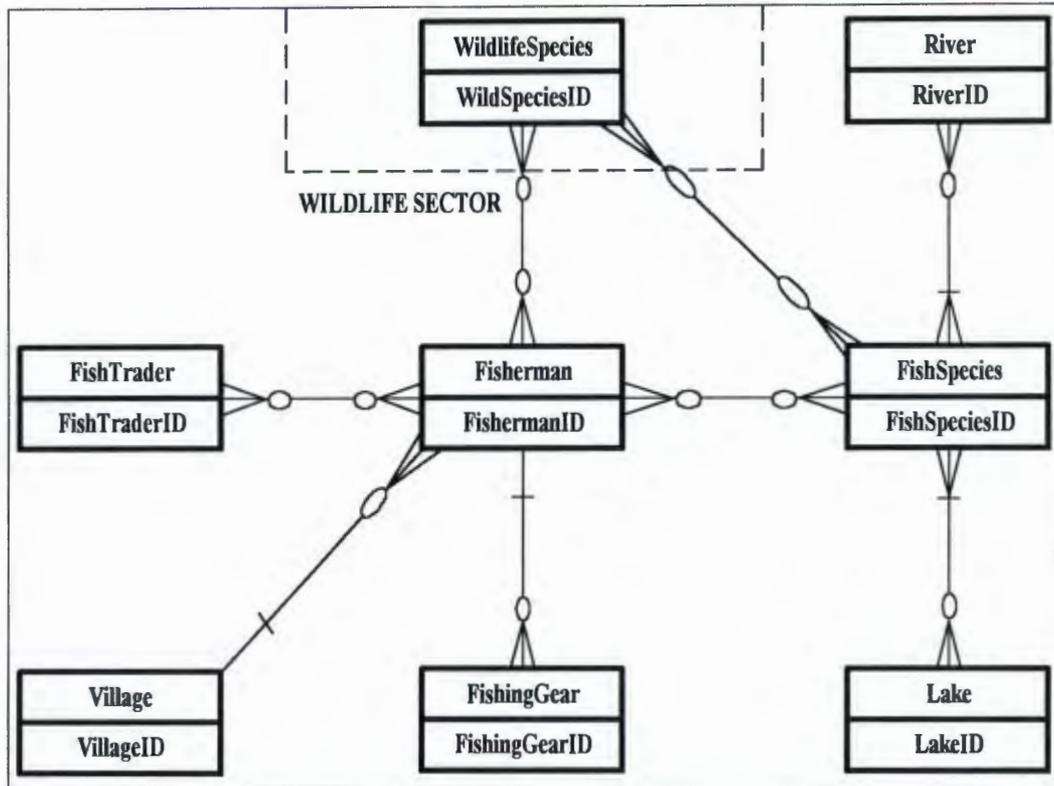


Figure 6.2 Fisheries Entity-Relationship Diagram

## 6.2.2 Beekeeping conceptual data model

The beekeeping conceptual data model consists of the enterprise rules that represent the interaction of the entities that make up the beekeeping sector. For instance, a beekeeper must operate at least one apiary while an apiary may belong to many beekeepers.

### **Beekeeper - Apiary**

A beekeeper must operate one or more apiaries

An apiary may belong to many beekeepers

### **Beekeeper - Honey Trader**

A beekeeper may sell honey to many honey traders

A honey trader may buy honey from many beekeepers

### **Apiary - Forest Reserve**

An apiary may be set up in a forest reserve

A forest reserve may have many apiaries

### **Apiary - Agriculture Area**

An apiary may be set up in an agricultural area

An agricultural area may have many apiaries

### **Apiary - Disease**

An apiary may be attacked by many diseases

A disease may affect many apiaries

### **Bee Species - Apiary**

A bee species may occupy many apiaries

An apiary must be occupied by a particular type of bees

### **Apiary Camp - Apiary**

An apiary (forest) camp may serve many apiaries

An apiary may be served by an apiary camp

### **Beekeeper - Wildlife Species**

A beekeeper may poach many wildlife animals

A wildlife species may be poached by many beekeepers

These enterprise rules are represented in the entity-relationship diagram in Figure 6.3 overleaf.

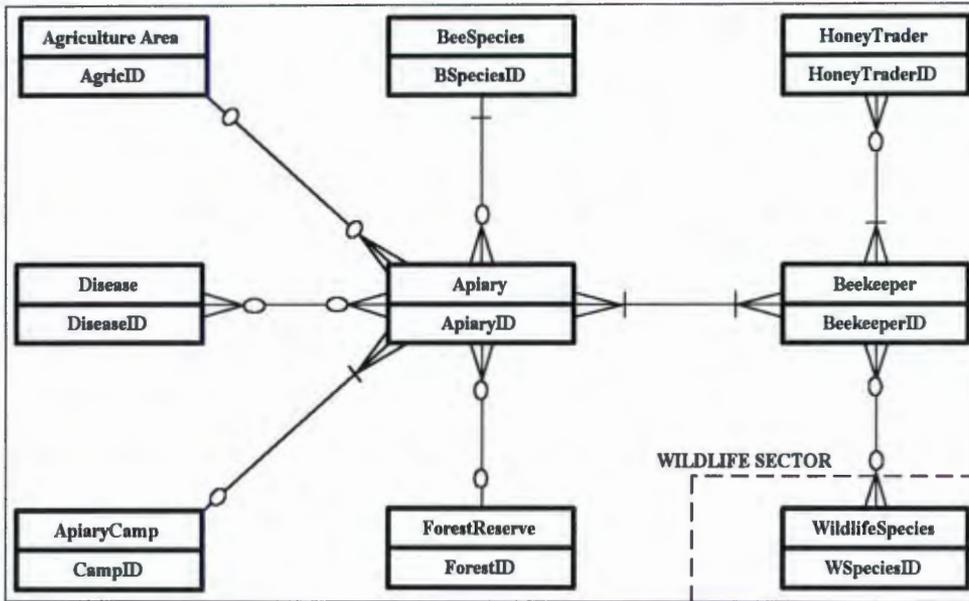


Figure 6.3 Apiculture Entity-Relationship Diagram

### 6.2.3 Wildlife Conceptual data model

#### Wildlife Enterprise Rules

##### **Local Hunter - Wildlife Species**

A local hunter must hunt one or more wildlife species

A wildlife species may be hunted by many local hunters

##### **Tourist Hunter - Wildlife Species**

A tourist hunter must hunt one or more wildlife species

A wildlife species may be hunted by many tourist hunters

##### **Local Hunter - Private License**

A local hunter may buy many hunting permits (private license)

A hunting licence must be held by one local hunter

##### **Tourist Hunter - Safari Operator**

A Safari operator may guide many tourist hunters

A tourist hunter must be accompanied by a Safari operator

##### **Safari Operator - Safari License**

A Safari operator may buy many hunting permits

A Safari licence belongs to a particular Safari operator

##### **Wildlife Species - Crop**

A wildlife species may destroy many crop types

A crop type may be destroyed by many wildlife species

##### **Crop - Agriculture Zone**

A crop type must grow in an agricultural zone

An agricultural zone may support many crop types

##### **Wildlife Species - Game Reserve**

A wildlife species may be found in many game reserves

A game reserve may support many wild life species

##### **Beekeeper - Wildlife Species**

A beekeeper may poach many wildlife animals

A wildlife species may be poached by many beekeepers

##### **Fisherman - Wildlife Species**

A fisherman may be attacked by many wildlife species

A wildlife species may attack many fisherman

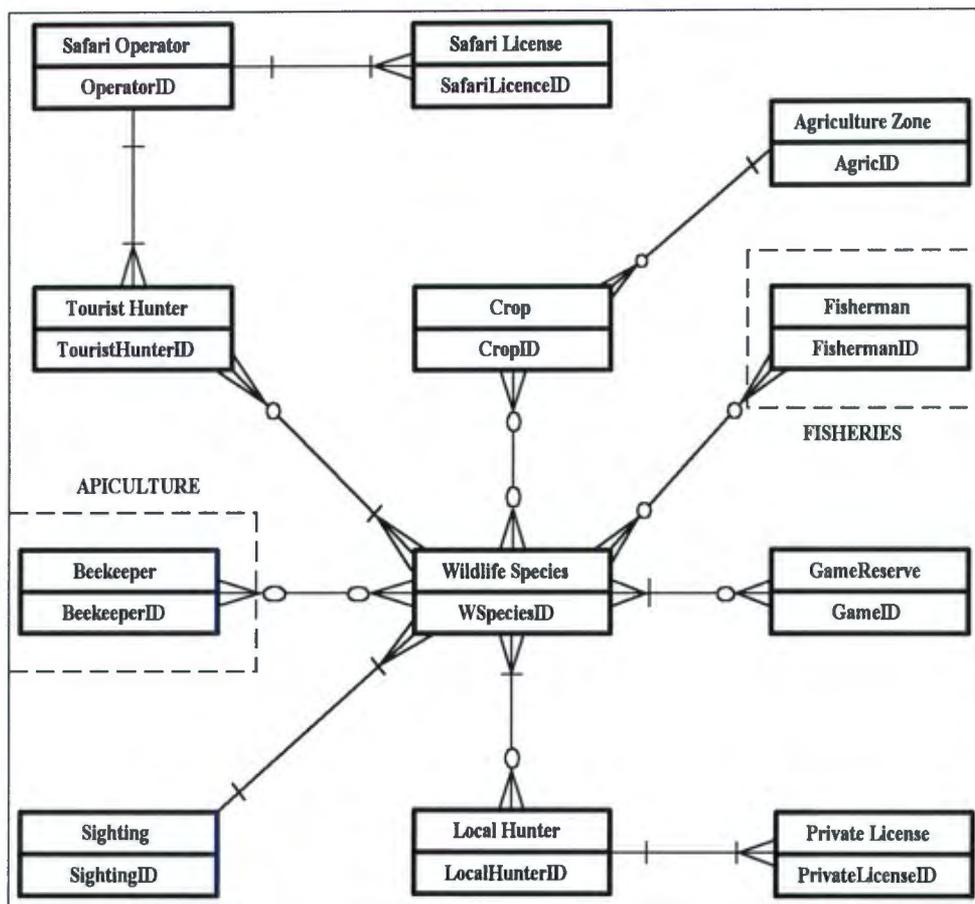


Figure 6.5 Wildlife Entity-Relationship Diagram

### 6.3 LRBIP-EIS logical data modelling

The LRBIP-EIS database will be based on the relational data model. A mapping process is applied to the conceptual data model so that it can be implemented in the relational data model. The mapping process involves the removal of those that are not compatible with the relational model (such as many-to-many relationships) and the derivation of the relations (skeleton tables) for the database. Skeleton tables do not show instances of the database tables. The relations show attributes that are associated with these tables including primary keys and any posted (foreign keys) that are used to uniquely identify them.

#### One-to-one (1:1) Relationships

If the relationship is mandatory on both sides, create a single relation with two candidate keys and choose one of the primary keys of the original keys to be the primary key of the

new relation, while the other is used as an alternate key. For a relationship that is mandatory on one side, create two relations and post the primary key of the optional entity in the relation of the mandatory entity

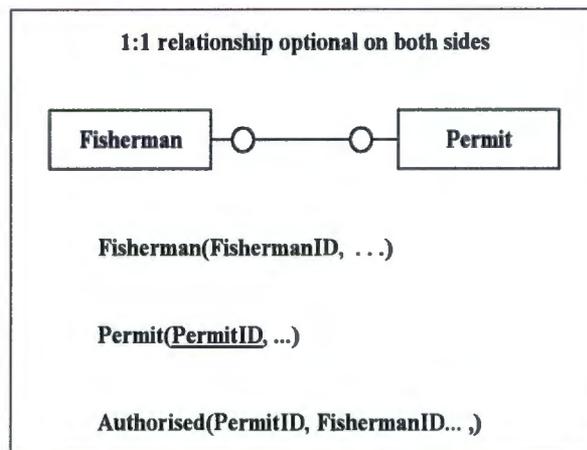
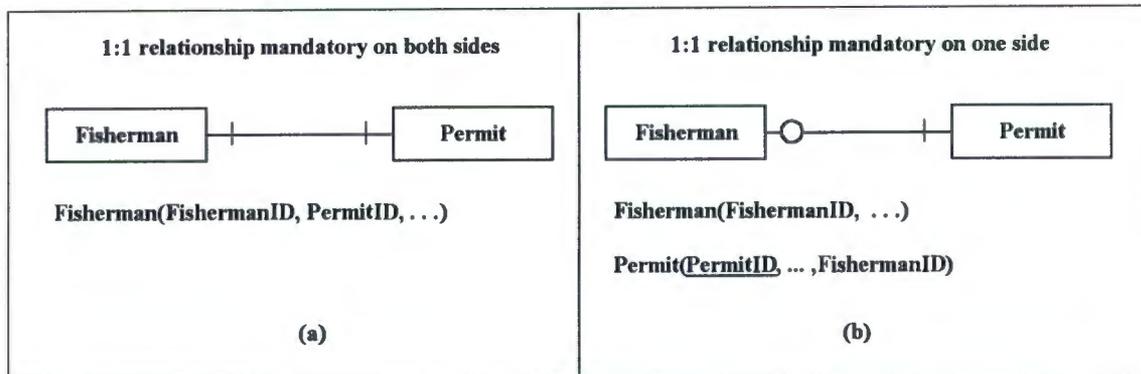


Figure 6.6 Mapping one-to-one relationships

If the relationship is optional on both sides, create three relations. The third relation shows the association between the entities and has two candidate keys (Figure 6.6 above).

### One-to-many (1: \*) Relationships

If the relationship is mandatory on the many side, create two relations and post the primary key of the optional side to the relation of the mandatory side. If the relationship is optional on the many entity, create a relation for each participating entity and a third relation for the association

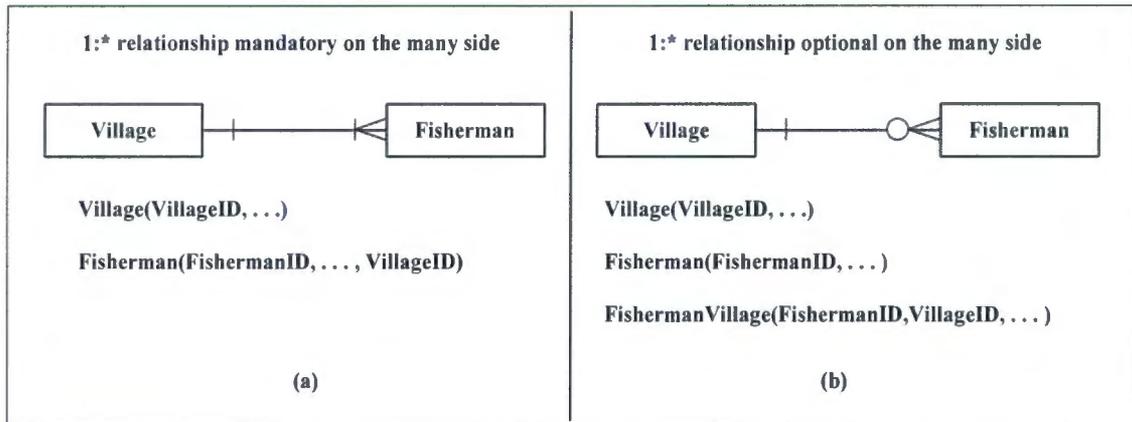


Figure 6.7 Mapping one-to-many relationships

### Many-to-many (\*: \*) Relationships

Create relations for the participating entities and a third one for the association with the primary keys of the participating entities forming a composite identifier.

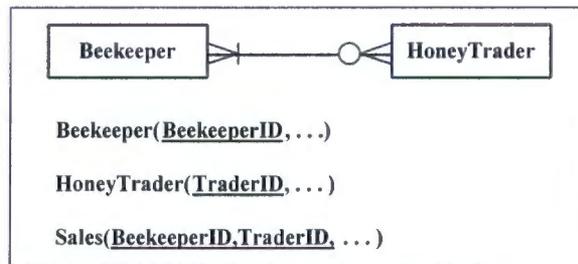


Figure 6.8 Mapping many-to-many relationships

#### 6.3.1 Fisheries logical data model

Relations represent the entities, attributes and entity relationships that have been identified during the conceptual and logical database design. Attributes are defined by the type of information that is required to held about an entity or relationship. The relations for entities and attributes are given by the name of the entity followed by its associated attributes in brackets. The logical data models for the wildlife (Section 6.3.2) and beekeeping (Section 6.3.3) sectors are derived in the same manner.

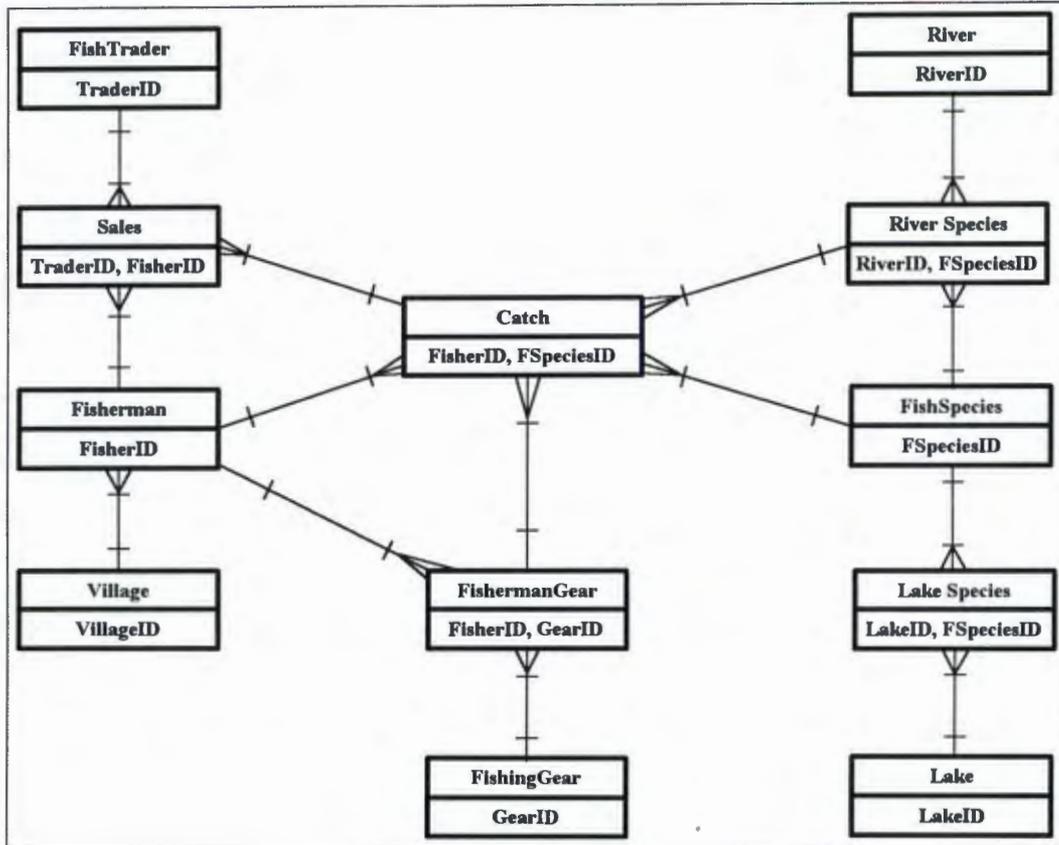


Figure 6.9 Fisheries Logical Data Model

**Fisheries Skeleton tables**

**Rivers** (RiverID, RiverName, Length)

**Lakes** (LakeID, LakeName)

**RiverSpecies** (RiverID, FSpeciesID )

**FishSpecies** ( FSpeciesID , SpeciesName, CommonName, LocalName, Family, Dimension, Occurrence, AquariumUse)

**RiverCatch** (FisherID, FSpeciesID, CatchDate , CatchSite, Weight, RiverID, GearID, TransportMode, Crew, FishingPeriod)

**Fisherman** ( FisherID, FirstName, Surname, Gender, DateOfBirth, MaritalStatus, Religion, VillageID )

**FishingGear** (GearID, GearType)

**FishermanGear** (FisherID, GearID , PurchaseDate)

**WildlifeSpecies** (WSpeciesID, SpeciesName)

**FishTrader** (TraderID, FirstName, Surname, Gender)

**Sales** (FisherID, TraderID , FSpeciesID, Quantity, Price, SaleDate, FishMarket)

**Village** (VillageID, Name, VillageHead, Population)

### 6.3.2 Beekeeping logical data model

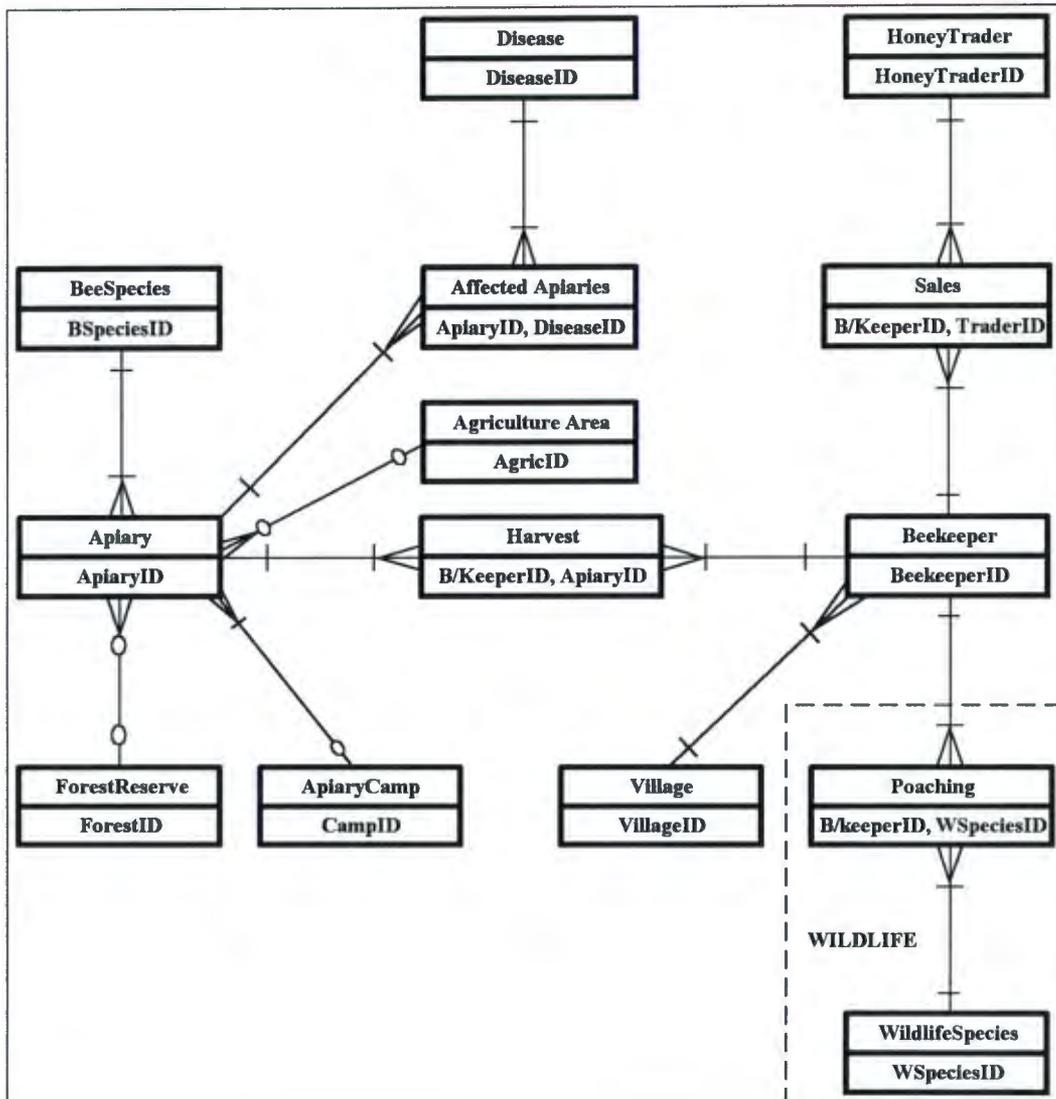


Figure 6.10 Apiculture Logical Data Model

#### **Beekeeping Skeleton tables**

**BeeSpecies** (BSpeciesID, Name)

**Apiary** (ApiaryID, BSpeciesID, CampID, Xcoord, Ycoord, TotalHives)

**ForestReserve** (ForestID, ForestName)

**ForestApiaries** (ApiaryID, ForestID)

**ApiaryCamp** (CampID, CampName)

**AgricultureZone** (AgricID, ZoneName)

**AgricApiaries** (AgricID, ApiaryID)

**Disease** (DiseaseID, DiseaseName)

**AffectedApiaries** (ApiaryID, DiseaseID, OutbreakDate)

**Harvest** (ApiaryID, BeekeeperID, Quantity, Date)

**Beekeeper** (BeekeeperID, FirstName, Surname, Gender, DateOfBirth, MaritalStatus, Religion, VillageID)

**Sales** (BeekeeperID, TraderID, Quantity, Price, SaleDate, Market)

**Village** (VillageID, Name, VillageHead, Population)

**HoneyTrader** (HoneyTraderID, FirstName, Surname, Gender)

**Poaching** (BeekeeperID, WSpeciesID, Date, Location)

**WildlifeSpecies** (WSpeciesID, Name)

### **6.3.3 Wildlife Logical Data Model**

#### **Wildlife Skeleton tables**

**WildlifeSpecies**(WSpeciesID, Name)

**Siting** (SitingID, WSpeciesID, DateSited, Xcoord, Ycoord)

**GameReserve** (GameID, GameName)

**LocalHunter** (LHunterID, Name, Surname, Address)

**LocalCatch** (LhunterID, WSpeciesID, PLicenseID, GameID, Date, Location (X,Y))

**PrivateLicence**(PlicenseID, LHunterID)

**Species-Reserve** (WSpeciesID, GameID)

**CropAttacks** (CropID, WSpeciesID)

**Crop** (CropID, AgricID, CropName)

**AgricultureZone**(AgricID, ZoneName)

**FishermanAttacks** (FishermanID, WSpeciesID, Date, Location [x ,y])

**Fisherman**(FishermanID, FirstName, Surname)

**FishSpecies** (FSpeciesID, Name)

**FishEaters**(WSpeciesID, FSpeciesID)

**TouristHunter** (THunterID, FirstName, Surname, Gender, OperatorID, SLicenceID)

**TouristCatch** (THunterID, WSpeciesID, GameID, Date, Location (X,Y))

**SafariOperator**(OperatorID, OperatorName, Address)

**SafariLicence**(SLicenceID, OperatorID, DateIssued, QuotaDetails)

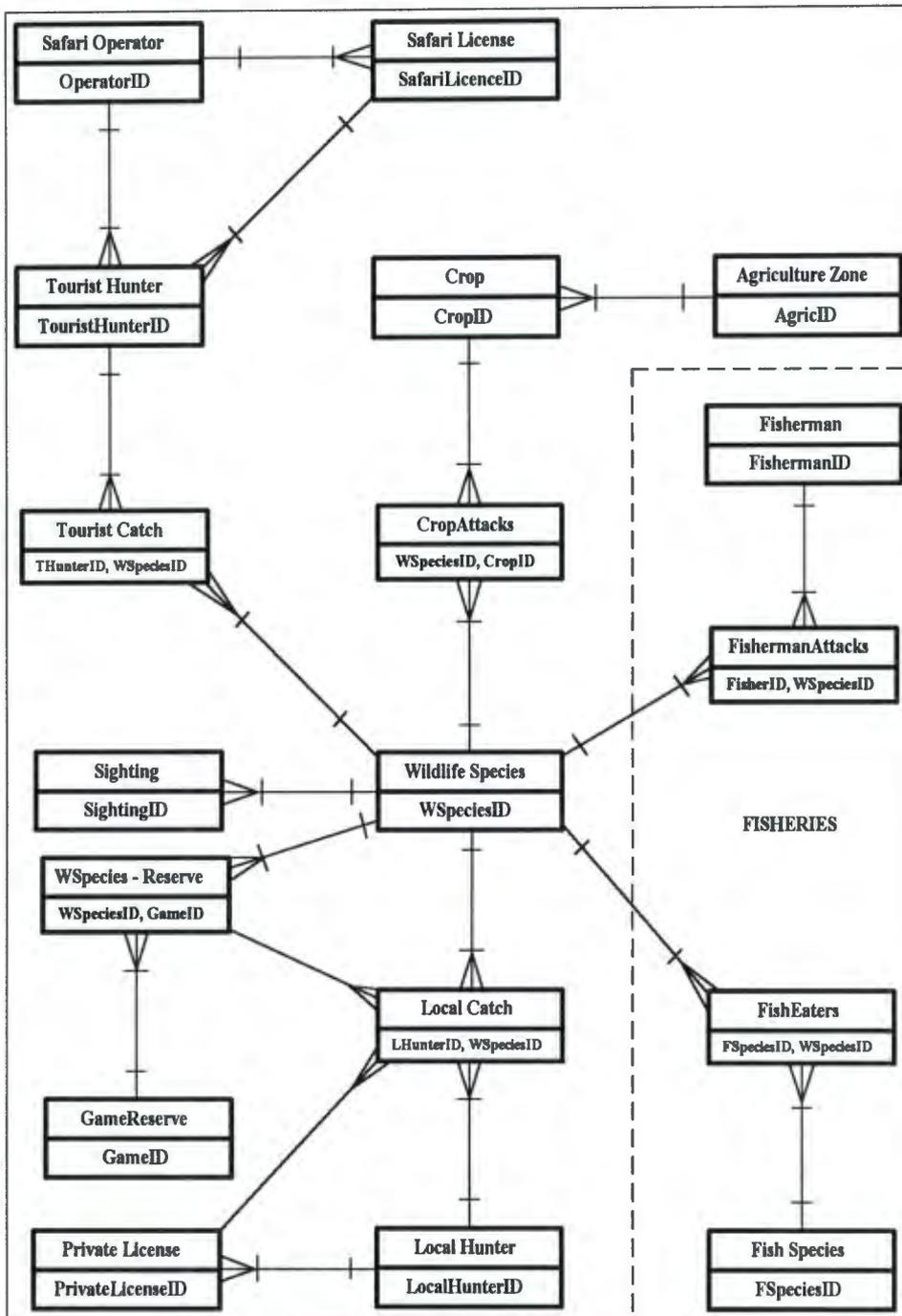


Figure 6.11 Wildlife Logical Data Model

#### **6.3.4 Validation of relations**

The relations in Section 6.3.1 to 6.3.3 above, need to be validated for anomalies using the process of normalization. The logical data models are also checked for fan and chasm traps before they are used for the design of the physical data model (Section 4.4 and 4.5).

#### **6.4 The Global Logical Data model**

The global logical data model integrates the local logical data models into a single model for the whole information system (Figure 6.12 overleaf).

#### **6.5 Physical Data Model**

The logical data model will be mapped onto a prototype physical data model in Microsoft Access Database Management System (a relational database environment). The process will involve designing of base relations, representation for derived data and defining some enterprise constraints. Entity and referential integrity constraints will be enforced. Environmental Systems Research Institute (ESRI) ArcView GIS and MapObjects packages will be used to handle the spatial component of the proposed EIS database.

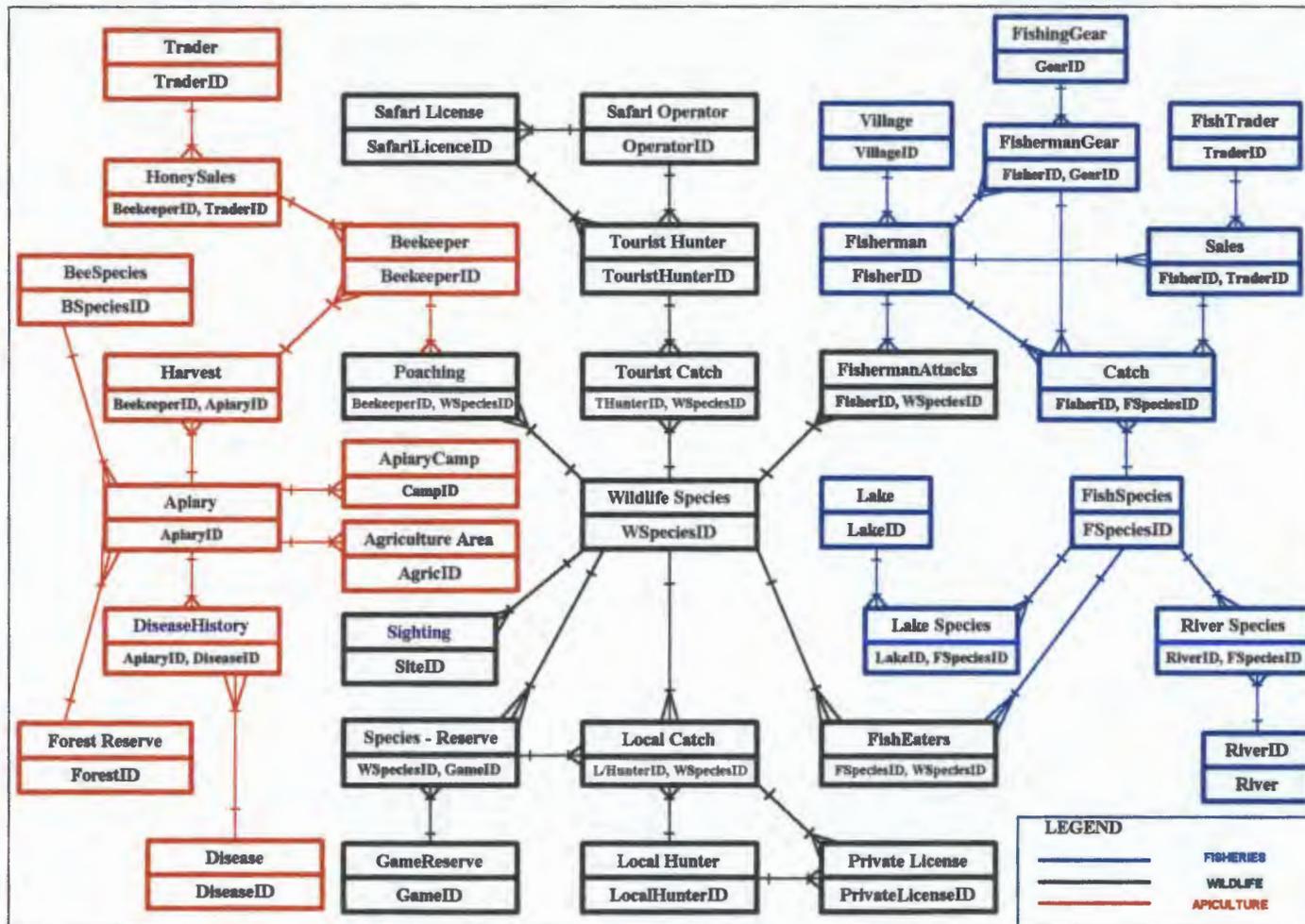


Figure 6.12 Global Logical Data Model

## **7 LRBIP-EIS PROTOTYPE DEVELOPMENT AND IMPLEMENTATION**

### **7.1 Introduction**

In the Software Development Methodology (**Section 3, Figure 3.1**), the implementation stage follows the systems design stage. Instead of developing full systems, prototypes are used to test new designs and ideas (Bhunu 1999). A prototype is a working model of the proposed system that is developed using sample data. Although the prototype is used for evaluation purposes, it also helps the developer to gain full experience with the kind of problems that can be expected in developing a full-blown system. The full system is developed after the evaluation of the prototype and adjusting the model designs if necessary. Due to time constraints, a prototype will only be developed for the fisheries sector.

The objectives of the fisheries database prototype are to provide:

- Data capture and viewing
- Data updating
- Querying and reporting textual and spatial data

### **7.2 Prototyping Scheme**

Figure 7.1 below shows the proposed scheme for the LRBIP-EIS database. The database has both spatial and non-spatial components. The figure shows how users will interact with the system as well as the interaction of the different softwares used. A graphical user interface (GUI) will be developed using Microsoft's Visual Basic (VB) environment and the Microsoft Access DBMS will provide the database engine for non-spatial data component of the system. ESRI ArcView GIS and MapObjects packages will be used to handle the spatial component of the system. The user will be able to capture, update and query the database through graphical user interfaces (GUIs). The users include district officers responsible for each environmental component, system administrators, regional and national administrative secretaries as well as internet browsers with interest in environmental information.

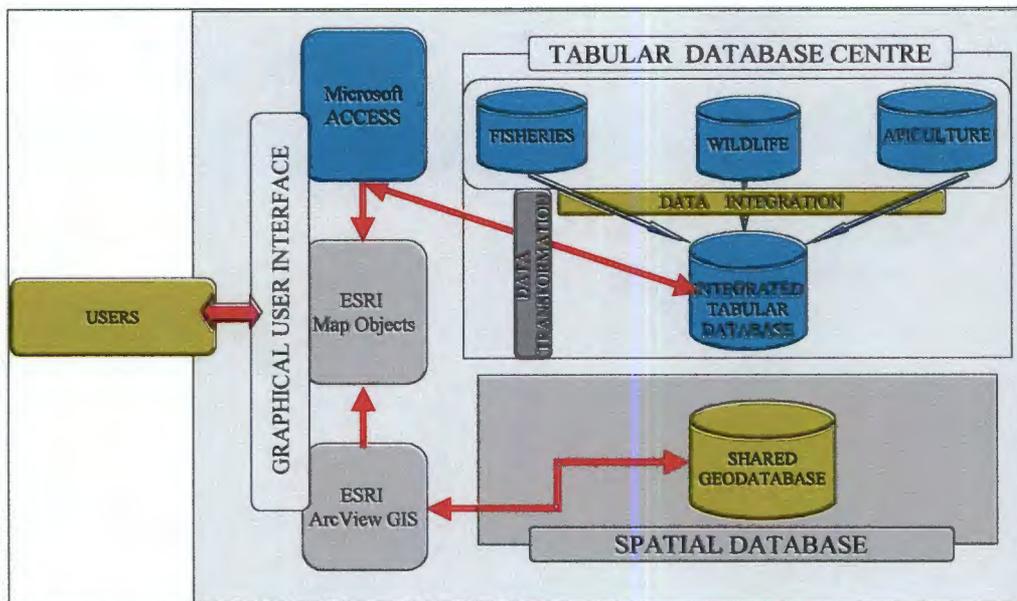


Figure 7.1 Proposed LRBIP-EIS database Scheme

### 7.3 Fisheries Test Database

In Section 6, three data models (fisheries, wildlife and apiculture) were integrated into a single logical data model that was developed for the LRBIP-EIS database. Developing a prototype to cover the designed model is beyond the time limitations of this research project. Therefore, the prototyping phase will concentrate on the Fisheries subsection of the proposed LRBIP-EIS database (see Figure 7.2 overleaf). The complete system can be developed on the basis of the evaluation of the Fisheries sector prototype. The following entities based on the logical data model in Section 6.3.1, represent some of the entities to be populated with sample data for testing the prototype:

- Fisherman
- Fish Species
- Lakes
- Rivers
- Fish Traders
- River Catch

- Lake Catch
- Fish Sales
- Villages
- Fishing Gear

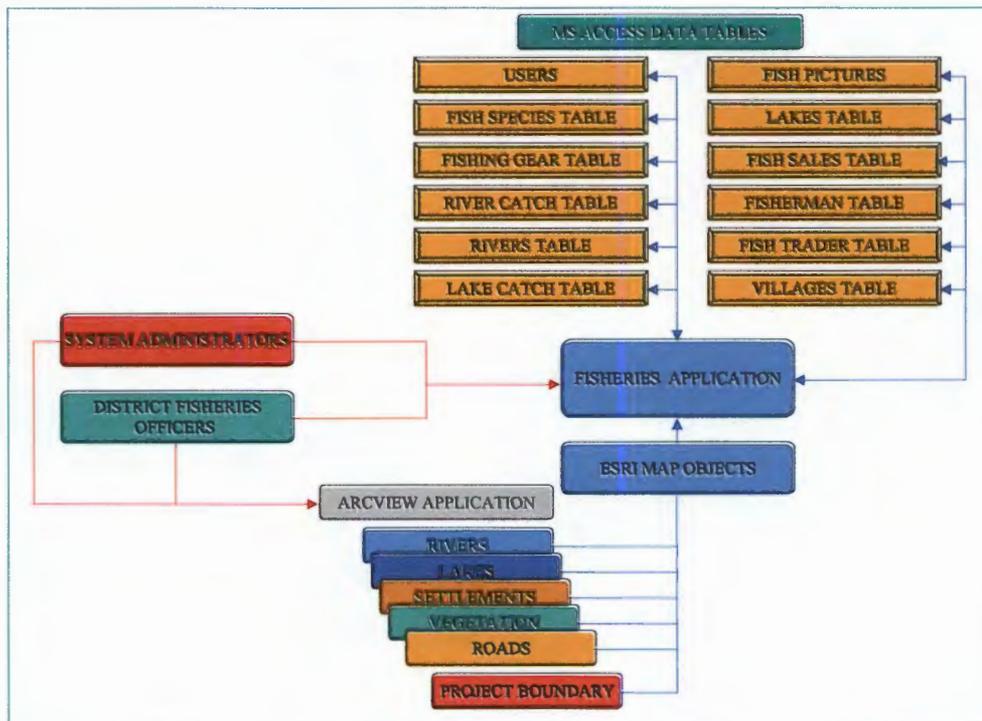


Figure 7.2 Fisheries Prototype Scheme

### 7.3.1 Non-spatial Database

The non-spatial component of the Fisheries prototype consist of two software environments, Microsoft (MS) Access and Visual Basic. The MS Access environment provides the platform for storing data and ensuring entity and referential integrity constraints are enforced. This is achieved by creating database tables from the skeleton tables (6.3.1) and then defining the relationships between the entities as shown in figure 7.3 overleaf. Appendix B shows samples of the *Fisherman* and *FishSpecies* tables and their associated data dictionaries.

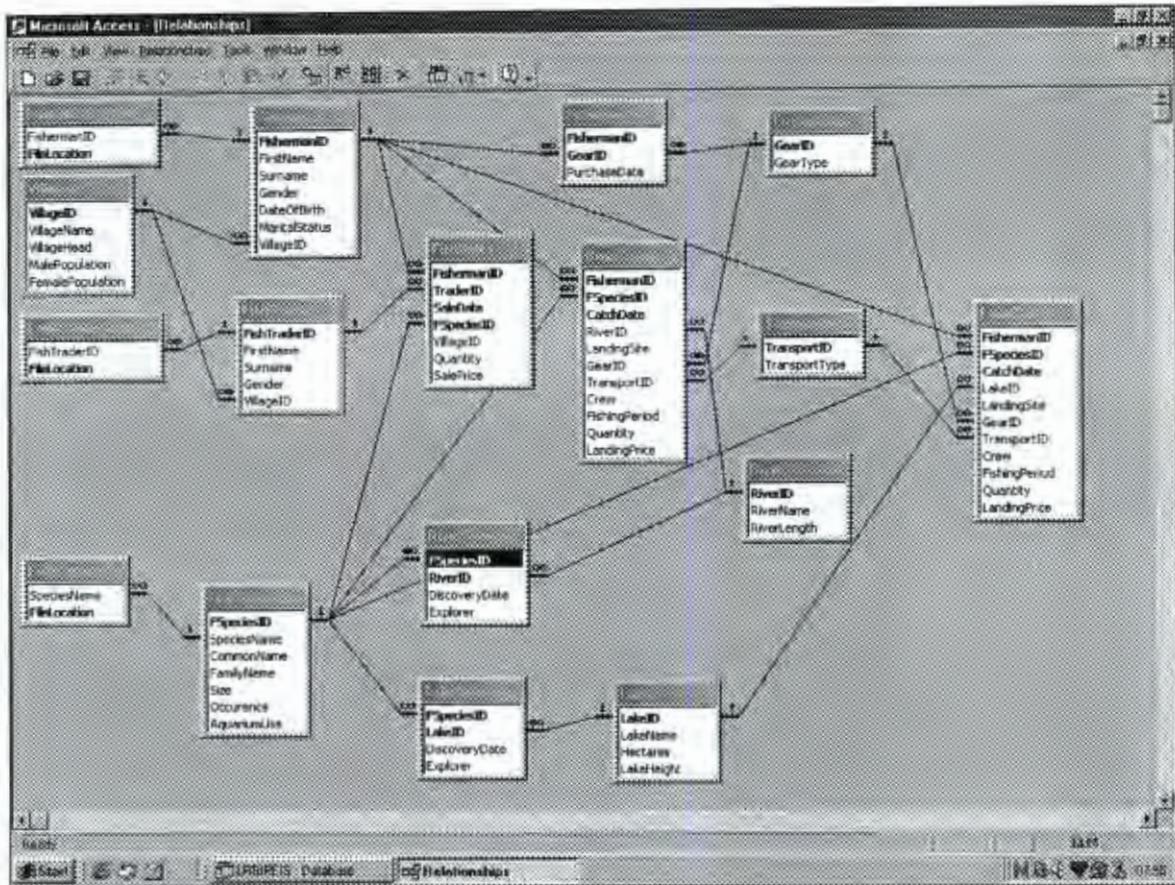


Figure 7.3 Microsoft Access Entity Relationships

The Visual Basic (VB) component is used for creating graphical user interfaces that display data from the MS Access database. This component is designed to enable users to capture, update and query the MS Access database. The VB component is based on the principles of multitier database application design (Bradley et al, 2001). In a multitier database, the application is broken down into independent components that should continue to do their job if the user interface changes (front end) or if the database changes (back end). Visual Basic programming was used extensively to build this component (see *Fisheries Application CD* in Appendix C). The reader is recommended to consult relevant literature on Visual Basic programming, if necessary. A three-tier model was adopted for prototyping the Fisheries component of the LRBIP-EIS database. In the model, each entity is implemented using three layers: User Services, Business Services and the Data Services.

- **User Services**

The User Services provide the user interface (e.g. forms, controls, and menus). This tier is made up of the various forms and the code that a user interacts with when dealing with a particular entity such as the form *frmFisherman* below. See the Fisheries.vbp project on the CD in Appendix C for more forms and their related code.

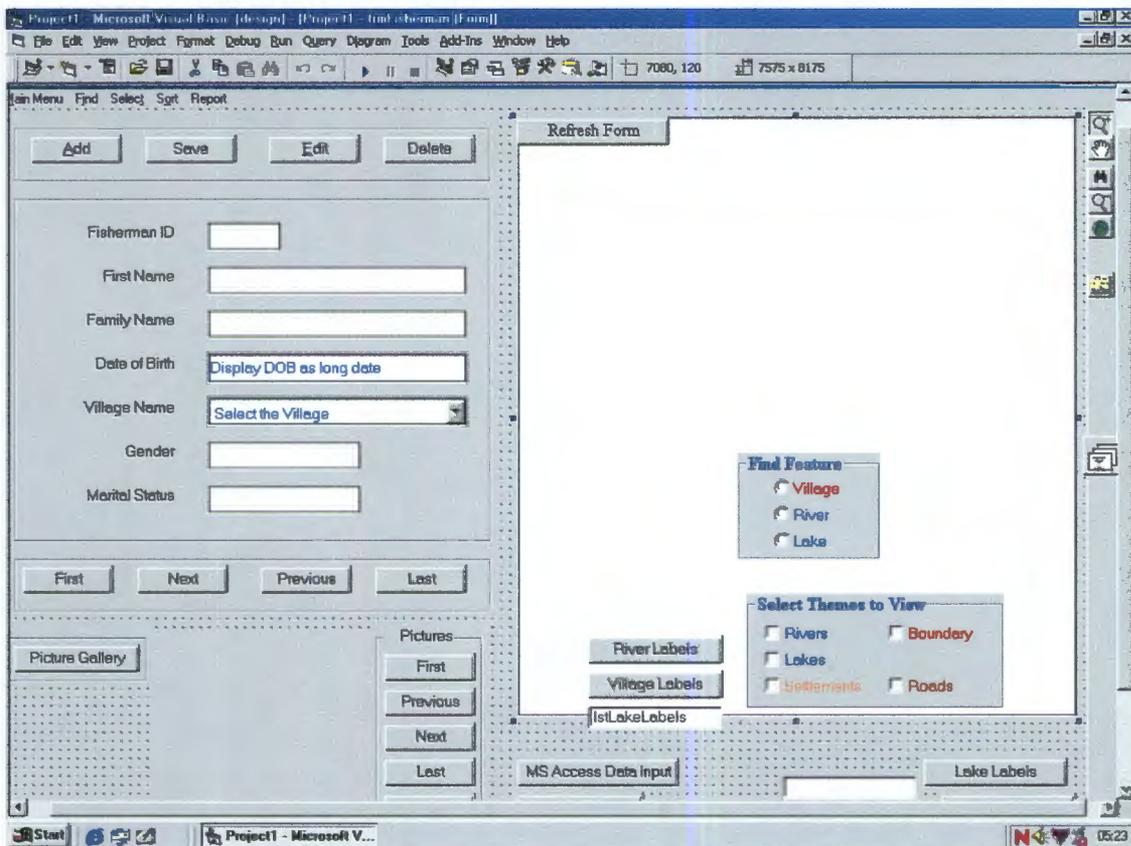


Figure 7.4 *frmFisherman* form during design

The User Services is used to display all messages to the user and maintains all user interface controls (e.g. whether the controls are visible or if the user can interact with a particular control).

- **Business Services**

The Business Services tier holds properties for each data field and performs all the validation, formatting (to comply with enterprise rules) and calculations on the data fields. This tier receives requests from User Services tier and calls the properties and methods of the Data Services. The Business Services tier is both a data consumer (from the Data Services tier) and a data provider (to the User Services tier). The *Property Let* procedures can perform any validation or formatting required for data input while the *Property Gets* provide the property values to the form. The Business Services uses a *BindingCollection* to bind properties to fields in a database table (See *CBSFisherman.cls* on CD in Appendix C for an example). If errors occur in the Business Services, they are raised and displayed in the User Services tier.

- **Data Services**

The Data Services tier interacts with the DBMS to store and retrieve data from the MS Access database. The Data Services tier handles data manipulation routines such as Add, Delete, Update, Sort, Find and Filter on a database table (recordset) that is currently active (see *CDSFisherman.cls* on CD in Appendix C). The Data Services tier is a data provider for the Business Services tier. Any errors that occur in the Data Services tier are reported in the User Services.

The major advantage of using multiple tiers is that a portion of the application can be changed without affecting all the code. For instance, migrating to a different database management system (such as Access Jet, SQL Server, Oracle or Sybase) only requires changes to the Data Services tier. If users require a different GUI, then changes to the users GUI should only affect the User Services tier.

### **7.3.2 Spatial Database**

The spatial data (such as rivers, lakes, villages, roads, vegetation) was provided by data providers such as TANRIC and Infobridge. The ERSI ArcView GIS package was used to

edit the spatial data, and to create only shapefiles for the case study area for each of the spatial components. Shapefiles are a simple, non-topological format for storing the geometric location and attribute information of geographic features (ESRI ArcView GIS 3.2 online help). The prototype allows restricted access to the ArcView GIS shapefile data to minimise the chances of tampering with this data by inexperienced users. The shapefile data is displayed in the GUIs via ESRI's MapObjects package. The MapObjects package is programmable in the Visual Basic environment (See *frmRiverCatch.frm* on CD in Appendix C). This functionality links textual data and spatial data allowing textual and spatial queries to be related. For example, selecting a fisherman from a MS Access table causes the fisherman's village to be highlighted in the mapping component (See Figure 7.8 in section 7.4.3).

## **7.4 Fisheries Prototype User Documentation**

### **7.4.1 Setting Up LRBIP-EIS prototype**

The fisheries prototype database has been packaged into a program file. The user should follow the instructions on set up CD in Appendix C. The *ReadMe.doc* file (on the same CD) contains detailed information on the hardware and software configuration that is required to set up and run the application successfully.

### **7.4.2 Navigating through the system**

Users can interact with the prototype using the various graphical user interfaces that have been designed to cater for the different user requirements.

#### **7.4.2.1 Logging In**

There are two ways of starting the prototype application. The first is to navigate to the program file containing the application using the Start icon on the status bar. The users can use a shortcut to the application from the desktop. The shortcut can be created after successfully setting up the prototype application. The GUI below pops up when the prototype begins to run.

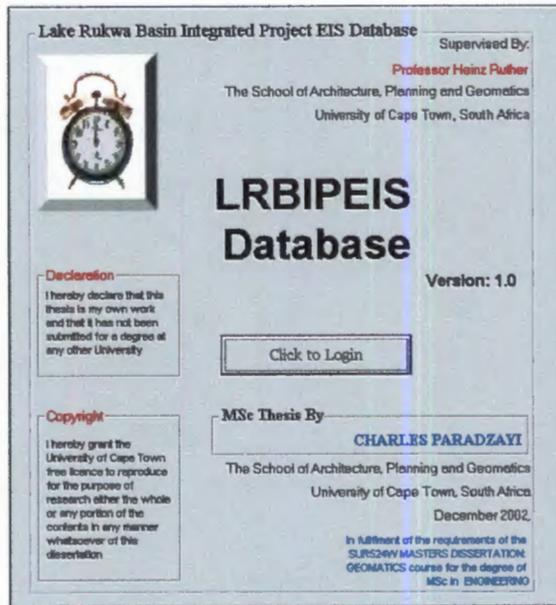


Figure 7.5 Initial graphical user interface

The user will be requested to login into the system. A valid username and password are required to gain access into the system. The system administrator is the only one allowed to add new users to the system.

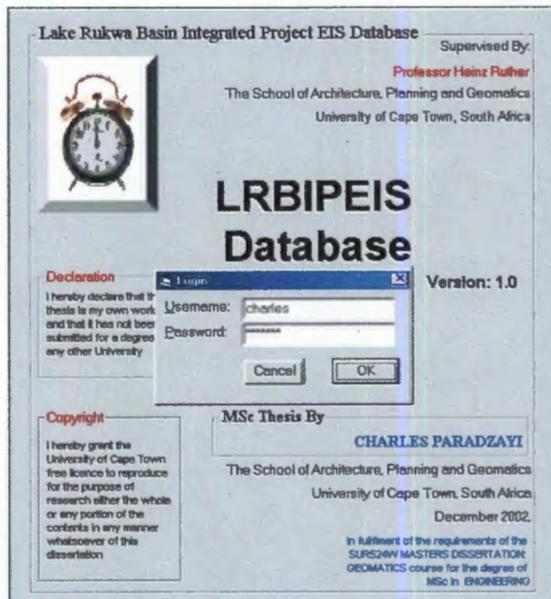


Figure 7.6 Login screen

When the user has successfully logged on to the system, then main menu is displayed, giving the user options to choose from as shown in figure 7.7 below. When the user selects an option, the system checks to see if that user is permitted to access the particular option. Access is granted only if a user has access rights otherwise a message will be displayed and control will be returned to the main menu. The access privileges are granted to each user when they register with the System Administrator. The Wildlife and Beekeeping sectors are still under construction.

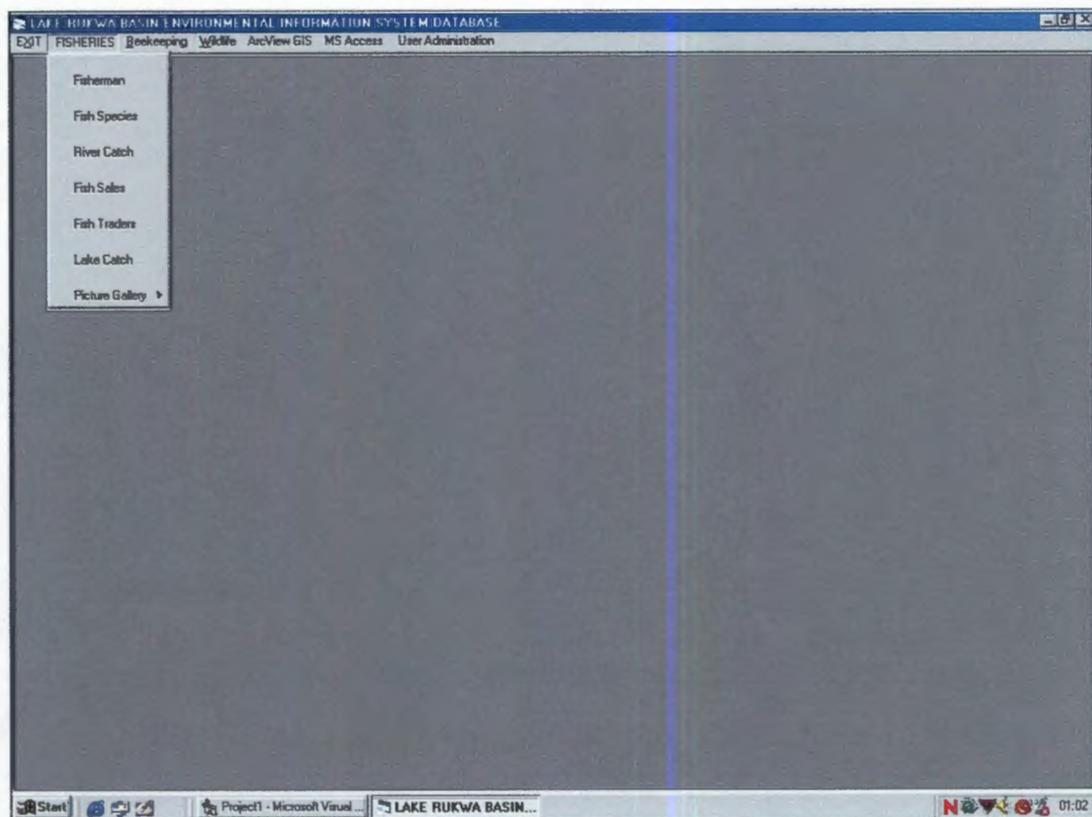


Figure 7.7 Main Menu

#### 7.4.3 Data Capture and Update

The Visual Basic forms are designed for the capture, viewing and updating of non-spatial data. Navigation buttons allow viewing of the first, next, previous or last record for the recordset being displayed. Data update (adding new record, deleting, editing or saving) is

achieved by clicking the appropriate control. The user should press the Tab key after entering attribute data or click on the next attribute control to proceed.

Spatial data updates are carried out in the ArcView GIS environment. Access to this package is restricted by means of access privileges that are granted to users as they register with the System Administrators. Detailed description of the operational nature of the GIS package is not presented here because it is expected that users with access rights to this package are experienced GIS users who can easily manoeuvre and update spatial data.

The MapObjects component has been programmed to give limited map viewing functionality. These include zooming, panning and spatial search options. A generalised display (with fewer themes) is shown when the user zooms to the full extent of the map display. This helps to reduce overcrowding in the display. Some themes can only be displayed when the user zooms in on the map display.

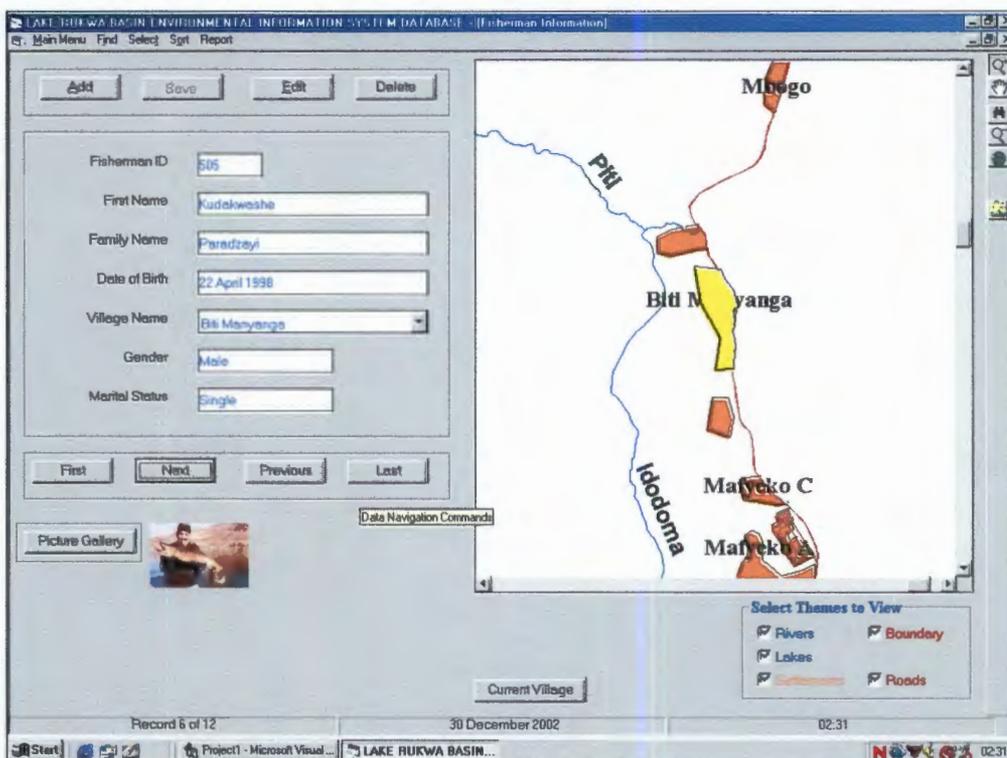


Figure 7. 8 Fisherman Data Capture, Update and Query interface



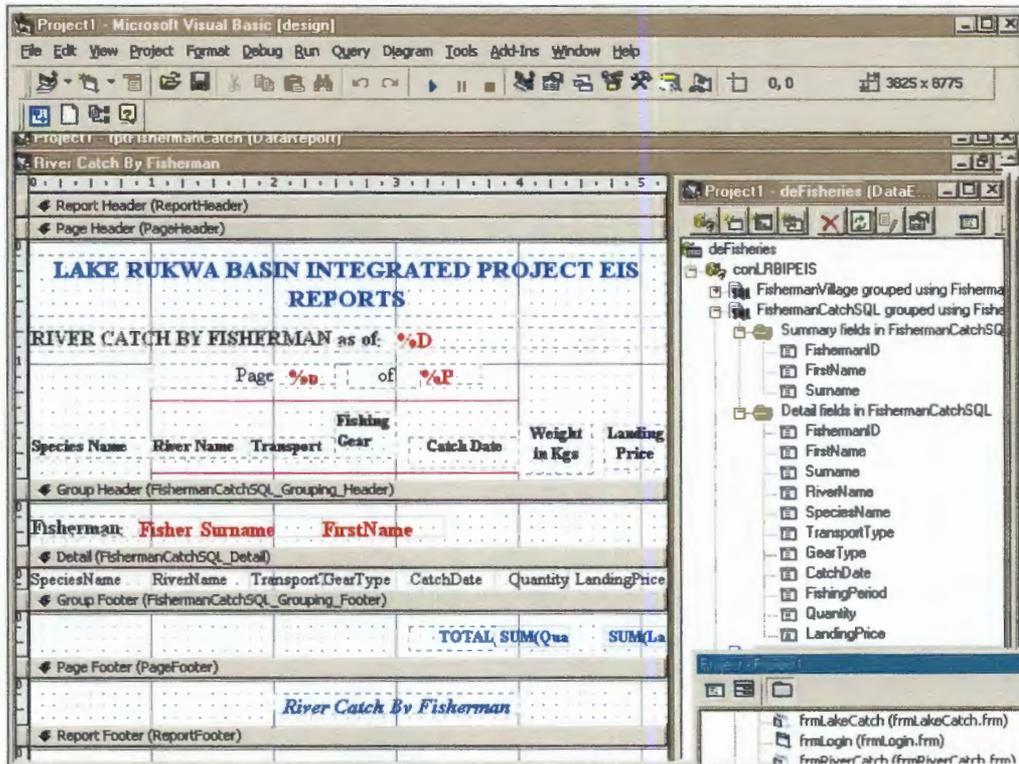


Figure 7.10 Sample Fisherman Catch Report Template

Figure 7.11 below shows part of a report on river fish catches by fishermen. The reports can be printed or exported in a variety of formats including one that is web-compatible.

River Catch By Fisherman

Zoom 100%

### LAKE RUKWA BASIN INTEGRATED PROJECT EIS REPORTS

RIVER CATCH BY FISHERMAN as of: **30 December 2002**

Page **1** of **1**

Species Name	River Name	Transport	Fishing Gear	Catch Date	Weight In Kgs	Landing Price
<b>Fisherman: 100 Makonese Charles</b>						
Protopterus	Apingile	Foot	Other	23 Nov 2002	234.0	TZ\$34.00
Oswochromis	Chalichambesi	Motor boat	Gillnet	10 Dec 2002	230.0	TZ\$2,300.00
<b>TOTAL</b>					<b>464.0</b>	<b>TZ\$2,334.00</b>
<b>Fisherman: 111 Ruhoto Charles</b>						
Momayrus	Chambwa	Foot	Gillnet 2	01 Jan 2002	125.0	TZ\$2,540.00
<b>TOTAL</b>					<b>125.0</b>	<b>TZ\$2,540.00</b>
<b>Fisherman: 112 Jamali Peter</b>						
Momayrus	Chambwa	Other	Gillnet 2.5	12 Dec 2001	233.0	TZ\$30.00
Hippopotamyrus	Chalichambesi	Cence	Beach Seine	12 Jan 2002	234.0	TZ\$1,236.00
<b>TOTAL</b>					<b>467.0</b>	<b>TZ\$1,266.00</b>
<b>Fisherman: 505 paradzayi kudakwasha</b>						
Haplochromis	Chaja	Foot	Beach Seine	20 Feb 2002	12.0	TZ\$3.00
Momayrus	Chipola	Cence	Long Line	18 Dec 2002	230.0	TZ\$2,500.00
Karis rukwensis	Idodoma	Foot	Long Line	22 Dec 2002	234.0	TZ\$3,000.00
Polypirus	Chimaga	Foot	Gillnet	28 Dec 2002	36.0	TZ\$360.00
<b>TOTAL</b>					<b>572.0</b>	<b>TZ\$6,065.00</b>
<b>Fisherman: 506 Paradzayi Dumai</b>						

Page: 1/1

Start Project1 - Microsoft Visual... LAKE RUKWA BASIN... Unfilled - Pair 02:51

Figure 7.11 River Catch By Fisherman Report

#### 7.4.6 Data Security

The application is only accessed via a username and password. A limited number of the users have access rights to both the MS Access and the ArcView GIS components of the database application. System administration has been limited to two people. The different sectoral officers are restricted only to their relevant sector, for example, fisheries officers can only access the fisheries component of the prototype.

#### 7.4.7 Limitations of the Prototype

The prototype has limited functionality with regard to some of the requirements outlined during the system design stage.

- An integrated database prototype could not be developed to cover the three sectors (fisheries, wildlife and apiculture) due to time constraints.

- Individual fisherman or fish trader reports were not developed.
- Weekly, monthly or yearly summaries were not generated for each fisherman or fish trader's activities
- The transfer of spatial attributes from the ArcView GIS environment to the MS Access environment require further investigation, so that changes in the ArcView environment become automatically available to the MS Access environment. At present, the user has to transfer the spatial attributes manually into the MS Access environment. This human element introduces redundancy that can lead to inconsistent data in the two environments.
- The Visual Basic component opens the MS Access database in an unsecured manner. This situation exposes the MS Access database to unauthorised access by any users who have Microsoft Access application. More research is required so that the VB component can open a password protected MS Access database.

## **8 EVALUATION AND RECOMMENDATIONS**

The research work involved the acquisition of skills from disciplines such as systems analysis, database design, computer programming and geographic information systems (GIS) in an attempt to meet the stated objectives (**Section 1.4**).

### **8.1 Evaluation**

The evaluation of this thesis is based on these objectives.

#### ***1. Review of the evolution of environmental information systems (EIS) in Africa***

The review showed that EIS development has been fragmented and has been constrained by technical, institutional and human resources limitations in countries or organisations that have attempted to implement environmental information systems.

#### ***2. Determine the information needs of the LRBIP-EIS database***

A thorough user needs assessment was conducted to identify the information requirements of the proposed database. Six environmental sectors (fisheries, beekeeping, wildlife, forestry, agriculture and mining) were identified as the core of the database for the case study area. A user needs collection guide was developed although it was not put into practice due to financial limitations to visit the intended users. Much of the information was gathered from literature reviews of the project documents as well as other database applications of a similar nature.

#### ***3. Conceptual, logical and physical data models for LRBIP-EIS database***

Due to time limitations on this research work, data modelling was confined to only three of the identified sectors i.e. fisheries, wildlife and beekeeping. A separate data model was developed for each sector before being integrated into a unified data model. This was the best approach because of the huge differences in individual sector information requirements. Areas of common interest between sectors were then identified and used to bridge the data models.

#### **4. *Develop and implement a prototype of the LRBIP-EIS database***

This component of the research posed the biggest challenge as it involved the integration of four computing environments to create the database application. The textual database was developed in MS Access environment because of its capability to enforce entity and referential integrity constraints on the data. ESRI ArcView GIS 3.2 software was chosen to maintain and update the spatial components. Visual Basic and MapObjects (Microsoft and ESRI products respectively) software was used for handling and visualisation textual and spatial data in an integrated environment. Visual Basic is very good for visual representation of textual data because customised graphical user interfaces (for data viewing, updating, querying and reporting) can be easily created. Spatial data can be viewed and queried in the same display as the textual data when Visual Basic is used in conjunction with the MapObjects component.

## **8.2 Recommendations**

A number of recommendations can be made from the analysis of this research work.

- The proposed user needs assessment guide (Appendix A) should be used to interview the district and regional officials who will use the database application. The result should be used to customise the prototype to meet the exact local demands of both textual and spatial information
- Prototypes should be developed for the wildlife and beekeeping components by customising the fisheries prototype. These prototypes should then be integrated with the fisheries to create an integrated working prototype which should be evaluated by the users. The users recommendations should be used to refine the prototype to produce the final LRBIP-EIS database application.
- Depending on the availability of reliable telecommunications and networking capabilities in the district offices, it is recommended that an intranet should be

developed for the Lake Rukwa Integrated Basin Project. Additional workstations, to complement the existing one, should be purchased for the individual sectors. The intranet should be designed to give sector users customised views of centrally held data. This will require more research into database security and the running of the database application on a shared client/server architecture (multi-user environment).

- River and lake catch locations should be fixed by GPS equipment for effective monitoring of fishing grounds. This information should alert authorities of areas that are being heavily exploited.

## REFERENCES

Atzeni P., Ceri S., Paraboschi S., and Torlone R., 1999, *Database Systems: concepts, languages & architectures*, McGraw-Hill Publishing Company

Bell R. V., Gerletti G., Soncini G., and Tosi G., 1995, *Mission Report, Conservation and use of environmental resources for the benefit of populations living in the area of Lake Rukwa, Lukwati and Muipa Forest Reserves, Tanzania*

Benyon D., 1990, *Information and Data Modelling*, Alfred Waller Ltd

Bhunu S. T., 1999, *Asset Management Information System Water Infrastructure in Harare, Zimbabwe*, MSc Thesis, ITC, Enschede, The Netherlands

Bradley J. C., and Millspaugh A. C., 2001, *Advanced programming Using Visual Basic 6.0*, McGraw-Hill Higher Education

Chiwandamira L., and Mbengo I., 1999, *Environmental Management, Module GED201*, Zimbabwe Open University

CIC Report, author unknown, *Conservation and use of environmental resources for the benefit of the populations living in the area of the basin of Lake Rukwa and in the protection belt east of the "Game and Forest Reserves" of Lukwati and Muipa - Tanzania*, undated

Connolly T., and Begg C., 2002, *Database Systems A Practical Approach to Design, Implementation, and Management*, 3<sup>rd</sup> Edition, Pearson Education Limited

Crispoldi A., (2001), *Sustainable statistical development for fisheries: role of the international institutions*, Conference on Agricultural and Environmental Statistical Application in Rome (CAESAR): Roma: 5-7 Giugno 2001, [www.istat.it/caesar/caesar2/inverventi/5giugno/21-Crispoldi.pdf](http://www.istat.it/caesar/caesar2/inverventi/5giugno/21-Crispoldi.pdf) (accessed 06/06/2002)

Dahlberg R., E., McLaughlin J., D., and Niemann B., J. (Editors) (1989), *Developments in Land Information Management*, Institute of Land Information, Washington D.C.

Davenport T., 2000, *The Lake Rukwa Basin Integrated Project, Technical Report for Centro Internazionale Crocevia CIC* – Roma, Sept 1999 – May 2000

Debriefing Note, Centro Internazionale Crocevia, 21 February 1999, unpublished.

EIS News, 1999. Ghana – Country at a Glance. [http://www.grida.no/eis-ssa/eisnews/1\\_99/cag.htm](http://www.grida.no/eis-ssa/eisnews/1_99/cag.htm) (accessed 23/01/2002).

EIS-SSA document, 2001. About the Program on Environmental Information Systems in Sub-Saharan Africa.

<http://www.grida.no/eis-ssa/about/about.htm>, (accessed 27/10/2001).

EIS-Uganda, 2002. Best Practices of EIS - The case of Uganda.  
[http://www.eis-africa.org/Uganda\\_1.html](http://www.eis-africa.org/Uganda_1.html) (accessed 11/01/2002).

GIS Development Guidelines, 2001, <http://www.sara.nysed.gov/pubs/gis/needs.htm>  
(accessed 15/10/2001)

Goodwin D., and Timberlake J., editors, 1993, *LIS/GIS – Getting it started* - Proceedings of the 4<sup>th</sup> seminar on Land and GIS in Zimbabwe, Harare

Harrington J. L., 2000, *Object-Oriented Database Design Clearly Explained*, Morgan Kaufmann Publishers

Hassan S. N., and Ndibalema V. G., 2001 *Apiculture: A Case Study of Lake Rukwa Basin, Tanzania*

Hawryszkiewicz I. T., 1991, *Introduction to systems analysis and design*, 2<sup>nd</sup> edition, Prentice Hall

Kalensky Z. D., and Latham J. S., 1998, *The Establishment of Environmental Information Systems in Developing Countries*  
[http://cgdi.gc.ca/sdi98/geomatica\\_papers/kalensky.pdf](http://cgdi.gc.ca/sdi98/geomatica_papers/kalensky.pdf) (accessed 23/06/2002)

Lake Tanganyika Biodiversity Program, 2002. [www.ltbp.org](http://www.ltbp.org) (accessed 15/01/2002).

Laurini R., & Thompson D., 1992, *Fundamentals of Spatial Information Systems*, Academic Press Limited

Lyytinen K., 1991, Penetration of Information Technology in Organisations.

<http://iris.informatik.gu.se/sjis/vol3/kalle.shtml> (accessed 10/10/2001).

MacDevette D.R., Forsyth G. G., and Jairoso Y., *THE USE OF PERSONAL COMPUTER-BASED GIS FOR NATURAL RESOURCE MANAGEMENT, 1995*:  
<http://easd.org.za/Publicat/pcgis95.htm>, (accessed 15/05/2001)

Mbassa G. K, and Mgongo F. O. K., 2001, *Agriculture, Tobacco Farming and Pastoralism in Makongolosi area of Chunya District*

Michener W. K., Brunt J. W, and Stafford S. G., 1994, *Environmental Information Management and Analysis: Ecosystem to Global Scales*, Taylor & Francis

Mukwada G., 2000, *Natural Resource Conservation and Management, Module GED404*, Zimbabwe, The Zimbabwe Open University

Musendo Z., 2001, *Land policies help impoverish indigenous communities*, Tanzania,  
[http://www.peacelink.it/afrinews/66\\_issue/p3.html](http://www.peacelink.it/afrinews/66_issue/p3.html) (accessed 22/10/02)

NEMC, 1998, *National Environmental Information Management Framework*,  
<http://easd.org.za/Eis/Tanzania/report.htm>, (accessed 13/04/2002)

NOAA, 2002, *Proposed Implementation of A Fishing Vessel Registration and Fisheries Information System*, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, [www.nmfs.noaa.gov/sfa/401.pdf](http://www.nmfs.noaa.gov/sfa/401.pdf) (accessed 12/06/2002)

Pratt P. J. and Adamski J.J., 2000, *Concepts of Database Management*, 3<sup>rd</sup> Edition, Thomson Learning

Prévost, Y. and Gilruth, P., 1999. *Environmental Information Systems in Sub-Saharan Africa: From innovation to management.*

<http://www.worldbank.org/afr/findings/english/find128.htm> (accessed 10/11/2001).

Rainer R.K., 2002, *Managing organizational data and information*, [www.auburn.edu/~rainerk/data\\_web.html](http://www.auburn.edu/~rainerk/data_web.html) (accessed 15/07/2002)

RFIS (Regional Fisheries Information System), 2001, *IT Assessments of the Coastal SADC States*, [www.schoemans.com.na/sadc/pdf/rfis\\_report%205%20IT.pdf](http://www.schoemans.com.na/sadc/pdf/rfis_report%205%20IT.pdf) (accessed 20/05/2002)

Rüther, H., 2001, *LRBIP Technical Report of the Technical Scientific Coordinator, Cape Town*, unpublished

Rüther, H., 2001, *Report on the Scientific Coordinator's first visit to the LRBIP Project, Mbeya*

Rüther, H., 2001. *EIS Education in Africa – The Geomatics Perspective.*

<http://www.ddl.org/figtree/pub/proceedings/nairobi/ruther-PS3-1.pdf> (accessed 22/01/2002).

Rüther H., 2001, *Lake Rukwa Basin Integrated Project Technical Report of the Scientific Coordinator*

Ssentongo G. W and Jihuliya J., 2000, Report on the Tanzania Fisheries Sector Review, FAO, Harare, [www.tzonline.org/pdf/reportonthetanzaniafisheriessectorreview.pdf](http://www.tzonline.org/pdf/reportonthetanzaniafisheriessectorreview.pdf), accessed 18/05/2002

Silberschatz A., Korth H., and Sudarshan S., 1997, *Database System Concepts*, McGraw-Hill

Smith R., 1999. A Consolidated Effort to sharing Metadata in South Africa. <http://www.grida.no/eis-ssa/eisnews/1-00/09.htm> (accessed 22/01/2002).

UNEP document, 2002. Environment and Natural Resources Information Networking – Africa Region. <http://www.unep.org/unep/eia/ein/enrin/enafrika.htm>, (accessed 18/01/2002).

Wiggins L.L. and French S.P., 1991, *GIS assessing your needs and choosing a system*, Report No. 433, American Planning Association

Wiid S., 2000, *The Software Development of the Tenement Information Management System*, Final Year Thesis, University of Cape Town

WLIP, 1991. Wisconsin Land Information Program. Wisconsin Land Information Board, <http://www.doa.state.wi.us/olis/wlip/> (accessed 5/3/2002).

World Bank Group Report, 1999, Environmental Information Systems in Sub-Saharan Africa: From innovation to management, <http://www.worldbank.org/afr/findings/english/find128.htm>

wp\_wdbfits.htm, 2002, *Which Client/Server Database Technology Fits? Experts and Practitioners Describe Alternatives*, [www.pds-site.com/WhitePapers/wp\\_wdbfits.htm](http://www.pds-site.com/WhitePapers/wp_wdbfits.htm) (accessed 15/07/2002)

Yeates D., Shields M., and Helmy D., 1994. *Systems Analysis and Design*. Pitman Publishing, London

## **APPENDIX A**

### **USERS' REQUIREMENTS COLLECTION GUIDE FOR LRBIP-EIS**

This requirements collection guide is an adaptation of one proposed by Wiggins et al, undated, for the City of Somerville, Massachusetts. The major objective is to document the problems, constraints, opportunities, requirements and priorities of the enterprise and the users of the system.

#### **INTRODUCTION**

Interviewer introduces himself/herself, briefly defines some of the key terms, gives an overview of the information gathering process, and highlights the areas of where data is being sought such sectoral information, current and planned attribute database and spatial database activities, issues concerning computer usage, and the sharing of data between and across sectors.

#### **ORGANIZATIONAL FUNCTIONS**

The respondent will be asked to describe his or her job functions, the sectoral structure, its size and its main activities that define their view of the proposed system.

1. What are the main functions of your sector?
2. How many employees are in your sector?
3. What specifically, is your function within the sector?
4. How long have been in your current position?

#### **CURRENT AND PLANNED GEODATABASE ACTIVITIES**

This section identifies the current and planned sectoral activities concerning geographic data, the information flows within and across sectors, and problems related to the use and maintenance of existing maps, if any.

5. What types of maps does your sector use in its work? [Summarize]

6. Can you provide some detailed information about your most important types of maps?

**Name of Map or Map Type**

- Natural features
- Boundaries
- Settlements
- Roads, sewer, waterlines, power lines
- Scale, level of accuracy, consequences of errors
- Georeferencing
- Source – organisational (consultant, in-house, other department, census)
- Source – physical (aerial photos, survey, scanning and digitizing)
- Complications (change of scale, piecing together maps)
- Routine operational use (e.g. sending public notices, weekly, monthly, annual reports)
- Analyses and decision-making use (suitability analysis, locating facilities, monitoring spatial patterns, impact analysis)
- Other comments

- How current are the basic data?
- How frequently is this data updated?
- Whose responsibility is it to do the update?
- What are the formal and informal updating procedures?
- What are the sources of information for updating the database?

7. Regarding current problems and future plans

- Are there any particular maps at specific scales that you would like to have but do not currently have?
- Have you had any problems with your current mapping activities? (e.g. out-of-date information, hard to retrieve, etc)
- Describe the types of analysis capabilities using maps that you want developed and made available in your department?
- Are there any other staff members in this department that you think I might want to interview because of their responsibilities in this area? Can I have their names and phone numbers or email addresses?

## CURRENT AND PLANNED DATABASE ACTIVITIES

This section identifies current databases (manual and/or automated attribute databases) that are or could be linked to the geo-spatial information used by the organization.

8. Can you provide some detailed information about some of the most important types of databases in your sector?

### Name of Database or Database Type

- Size (approximate number of records, number of fields/items)
- Snapshot (current data) vs. historical (time series with each transaction recorded)
- Georeferencing (geo-coding methods)
- Source – organisational (consultant, in-house, other department, census)
- Source – physical (census, survey, scanning and digitizing)
- Routine operational use (e.g. sending public notices)
- Analyses and decision-making use (permit tracking, monthly summary, monitoring spatial patterns, etc)
- Other comments

- How current are the basic data?
- How frequently is this data updated?
- Whose responsibility is it to do the update?
- What are the formal and informal updating procedures?
- What are the sources of information for updating the database?
- Any departmental staff designated to update the databases (re: security of data)?

9. Regarding current problems and future plans

- Are there any databases currently being developed within your sector? If so, what are they?
- What additional database capability would you like to see in your sector?
- What frustrations do you have with your current database systems? (e.g. lack of data, inaccessible data, limited flexibility in report formats)

## **COMPUTER USE AND DATA SHARING ACROSS DEPARTMENTS**

10. Could you please briefly describe your current computer resources and use?
11. Which specific software packages does your sector use frequently? When did your sector first start using these packages?
12. Please tell me about other software packages that you use occasionally in your sector.
13. Are there hardware and/or software resources that you currently do not have access to that you feel would help you carry out your job more easily and/or effectively?
14. Currently, do you share data with any other sectors? If so, briefly describe the maps and databases involved and the reason for sharing. [If no, then what data do you foresee as shareable with other sectors?]
15. Do you experience (or can you foresee) any problems when sharing maps and databases with other sectors e.g. conflicts over reference points or geographic mismatches? Are these due to methods of survey or levels of accuracy?
16. Are there any particular concerns about data security and control over access and updating that might complicate data sharing?
17. Which models are being used currently for analysis in your sector and how can they be incorporated into the proposed environmental information system?
18. Do you have any other comments or concerns that you would like to share with me at this time?

## APPENDIX B

### SAMPLE DATABASE TABLES AND DATA DICTIONARIES

Fisherman table

The screenshot shows a Microsoft Access window titled "Microsoft Access - [Fisherman : Table]". The window displays a table with the following columns: FishermanID, FirstName, Surname, Gender, DateOfBirth, MaritalStatus, and VillageID. The data is as follows:

FishermanID	FirstName	Surname	Gender	DateOfBirth	MaritalStatus	VillageID
100	Charles	Makonese	Male	2/1/69	Married	1
111	Charles	Ruhoto	Male	1/2/72	Single	4
112	Peter	Jamali	Male	12/15/50	Widowed	26
123	Makonese	Tichaona	Male	12/12/72	Single	26
124	Tinashe	Paradzayi	Female	10/3/83	Single	4
125	Innocent	Hapazari	Male	9/28/65	Single	6
126	Stashia	Paradzayi	Female	8/28/70	Married	13
200	Nehemiah	Mavetera	Male	8/15/68	Married	7
305	chipo	Munhu	Female	9/9/78	Single	5
505	kudakwashe	paradzayi	Male	4/22/98	Single	1
506	Danai	Paradzayi	Female	2/18/01	Single	4
1000	7sazsa	Milnngn	Female	2/5/75	Married	26
0						0

At the bottom of the window, it says "Record: 11 of 12" and "National identity number of fisherman".

Fisherman Data Dictionary

The screenshot shows a Microsoft Access window titled "Microsoft Access - [Fisherman : Table]". The window displays a data dictionary for the Fisherman table with the following columns: Field Name, Data Type, and Description.

Field Name	Data Type	Description
FishermanID	Number	National identity number of fisherman
FirstName	Text	First name
Surname	Text	Last name
Gender	Text	Gender of fisher
DateOfBirth	Date/Time	Date when fishermen was born
MaritalStatus	Text	Marital status of fisher
VillageID	Number	(Foreign key) VillageID references village from Village table

## Fish Species Table

FSpeciesID	SpeciesName	CommonName	FamilyName	Size	Rare or Comm	Aquarium Use?
1	Protopterus annectens brieni		Protopteridae	Big	Rare	No
7	Polypterus ornatipinnis		Polypteridae	Big	R	Yes
8	Hippopotamyrus discorhynchus		Mormyridae	M		
9	Marcusenius macrolepidotus		Mormyridae	M		N
10	Mormyrus longirostris		Mormyridae	B	C	
11	Petrocephalus catostoma		Mormyridae	S	R	N
12	Kneria rukwaensis		Kneriidae	S	R	N
13	Barbus apleurogramma		Cyprinidae	S	C	
14	Barbus brachygramma	Linespotted ufipa	Cyprinidae	S		
15	Barbus innocens	Inconspicuous barb	Cyprinidae	S		
16	Barbus jacksoni	Jackson's barb	Cyprinidae	S		
17	Barbus lineomaculatus "form 1"		Cyprinidae	S		Y
18	Barbus lineomaculatus "form 2"		Cyprinidae	S		
19	Barbus luikae		Cyprinidae	S		
20	Barbus oligogrammus		Cyprinidae	S		Y
21	Barbus olivaceus		Cyprinidae	S	R	
22	Barbus paludinosus		Cyprinidae	S	C	
23	Barbus pellegrini		Cyprinidae	S		
24	Barbus pseudotoppini		Cyprinidae	S		
25	Barbus radiatus		Cyprinidae	S	C	
26	Chelaethiops rukwaensis		Cyprinidae	S	C	
27	Labeo cylindricus		Cyprinidae	M	C	
28	Labeo fuelleborni		Cyprinidae	M	C	N
29	Mesobola spinifer		Cyprinidae	S		N
30	Raiemes moorii		Cyprinidae	M	C	Y
31	Brycinus imberi		Characidae	M	C	N
32	Brycinus andersoni		Characidae	S		N

## Fish Species Data Dictionary

Field Name	Data Type	Description
FSpeciesID	Number	unique fish species identifier
SpeciesName	Text	Scientific name of fish species
CommonName	Text	Commonly used name (English name) of the fish species
FamilyName	Text	Family name of fish
Size	Text	Qualitative size of the fish species
Occurrence	Text	Indication of whether the species is abundant or rare
AquariumUse	Text	Whether the fish species is good for aquarium use or not

**APPENDIX C**

**LRBIP-EIS SET UP CD WITH SOURCE CODE**

