Economic
Commission
for Africa

ECA/ISTD/GEO/2007/02E

Determination of Fundamental Datasets for Africa

Geoinformation in Socio-Economic Development

Ordering information

To order copies of *Geoinformation in Socio-Econmic Development: Determination of Fundamental Datasets for Africa* by the Economic Commission for Africa, please contact:

Publications
Economic Commission for Africa
P.O. Box 3001
Addis Ababa, Ethiopia

Tel: +251 11 544-9900
Fax: +251 11 551-4416
E-mail: ecainfo@uneca.org
Web: www.uneca.org

ECA/ISTD/GEO/2007/02E – Geoinformation in Socio-Economic Development: Determination of Fundamental Datasets for Africa

Prepared for:
The Working Group on Fundamental Data Sets
Of the Sub Committee on Geoinformation
Of the Committee on Development Information, Science and Technology
Of the Economic Commission for Africa

Completed Under Contract by:
The Human Sciences Research Council and EIS-AFRICA

Funded by:
Chief Directorate of Surveys and Mapping
Department of Land Affairs
South Africa

First Printed September 2007

Edited, designed and printed by the ECA Publications and Conference Management Section (PCMS).
Cover photo: USAID

Table of Contents

Acronyms

ADIE	l'Association pour le Developpement de l'Information Environnementale
ADS	Africa Data Sampler
AGRHYMET	Centre Regional de Formation et d'Application en Agrométéorologie et Hydrologie Opérationnelle
ANZLIC	Australia New Zealand Land Information Council
CEDARE	Centre for Environment and Development for the Arab Region and Europe
CODI	Committee on Development Information
DCW	Digital Chart of the World
DEM	Digital Elevation Model
DEWA	Division of Early Warning and Assessment
DGENV	Directorate General Environment
EA	East Africa
ECA	Economic Commission for Africa
EIS	Environmental Information Systems
ETeMII	European Territorial Management Information Infrastructure
EUROSTAT	Statistical Office of the European Communities
FAO	Food and Agriculture Organization
FGDC	Federal Geographic Data Committee
FIG	Fédération Internationale des Géomètres
GEO	Global Earth Observation
GI	Geographic Information
GIS	Geographic Information Systems
GLCN	Global Land Cover Network
GPS	Global Positioning System
GSDI	Global Spatial Data Infrastructure
GTOS	Global Terrestrial Observation System
HSRC	Human Sciences Research Council
ICA	International Cartographic Association
ISO	International Standards Organization

ISTD	ICT and Science & Technology Division
ITC	Information Technology Center
MAFA	Mapping Africa for Africa
MDG	Millenium Development Goals
NCS	National Conservation Strategy
NEAP	National Environmental Action Plan
NEPAD	New Economic Partnership for African Development
NGDI	National Geospatial Data Infrastructure
NSDI	National Spatial Data Infrastructure
NSIF	NATO Secondary Imagery Format
NTDB	National Topographic Data Base
ONC	Operational Navigation Chart
RCMRD	Regional Centre for Mapping of Resources for Development
RECTAS	Regional Centre for Training in Aerospace Surveys
RRSU	Regional Remote Sensing Unit
SADC	South African Development Community
SAHIMS	Southern Africa Humanitarian Information Management Network
SDI	Spatial Data Infrastructure
UN	United Nations
UNDP	United Nations Development Programme
UNECA	United Nations Economic Commission for Africa
UNEP	United Nations Environment Programme
UNGIWG	United Nations Geographic Information Working Group
US	United States
USA	United States of America
USGS	United States Geological Survey
VMAP	Vector Map
WHO	World Health Organization
WMO	World Meteorological Organization
ZAR	South African Rand

Determination of the Fundamental Geo-spatial Datasets for Africa[1]

Executive Summary

Recognizing that Africa's social and economic development is predicated on the careful assessment, identification, and operationalization of national development priorities, the role of high-quality data and information to support such informed decision-making readily comes to the fore. As geospatial data and geoinformation are widely accepted as essential components of the body of knowledge that informs national development strategy, then a pan-continental and common definition of what constitutes a minimally necessary core of geospatial data and information products is required. As such, the purpose of the present paper is to identify and enumerate these core, or fundamental, geospatial data sets to support Africa's development agenda.

A user needs analysis has been conducted involving acknowledged sub-regional, regional and global partners. The outcome of the desk study and interviews is that a set of fundamental data has been proposed.

In order to arrive at the fundamental data sets for the entire African continent, it was necessary to clearly define what constitutes fundamental data. However, the literature review shows that there is no universally accepted or unique definition of fundamental data sets. Using inputs from the various collaborators from the many international and African institutions and the literature reviewed, the study recommends that the following definition be adopted:

Fundamental data sets are the minimum primary sets of data that cannot be derived from other data sets, and that are required to spatially represent phenomena, objects, or themes important for the realisation of economic, social, and environmental benefits consistently across Africa at the local, national, sub-regional and regional levels.

Some of the universal and key criteria used in defining the fundamental data sets are apparent in the definition presented above. Other criteria and guidelines for identifying fundamental data sets should include complete coverage over the area of interest, it should be needed consistently, must have sufficient detail, and a diversity of users from different sectors must derive significant benefit from their use. Fundamental data sets should also have acceptable standards and validation processes that ensure consistency, reliability, quality, continuity and accuracy.

1 This report integrates inputs from international, regional, sub-regional and national contributors.

On the basis of the definition and criteria above the following data themes have been identified as constituting the fundamental data sets for Africa:

- Geodetic Control Network
- Imagery
- Hypsography
- Hydrography
- Boundaries
- Geographic names
- Land management units/areas
- Transportation
- Utilities and services
- Natural environment

These have been hierarchically classified into different levels, categories and themes based on all the inputs received on the criteria and definitions from the study. The levels reflect the relative and sequential importance of the different data sets in the development of a universal set of geo-spatial information for the African continent. The categorization reflects the functional uses of the fundamental data sets in terms of their use as a geographic reference frame, as base geography and as a geo-coding scheme needed to give non-spatial data a geographical reference.

The study also presented findings on what spatial features should form part of the fundamental data sets, what attributes should be associated with them, what level of detail the data sets should be developed at, what metadata should be developed and what requirements there are for the temporal updating of the fundamental data sets.

It was apparent from the study that fundamental data sets need to be identified within appropriate user-defined frameworks and in many instances, are defined by the mandate's of organizations. A key aspect of fundamental data sets is that they should be a reference frame, foundation or base for the development and integration of geo-spatial data sets within these frameworks at a national, sub-regional and regional level. For this to be accomplished it is necessary for the data to be available and widely accessible so that new geo-spatial data sets can be developed through the cooperation of users.

Information on issues presented in the executive summary and more, are described in greater detail in the body of the report.

1. Introduction

Characterized by data- and information-driven innovation, creativity, and flexibility, the contemporary global information economy is predicated on several reinforcing pillars: computerization and the intensive use of information; the codification of knowledge; the transformation of information products into commodities; and new ways of organizing work and production. A major imperative, then, for all countries to participate in the information economy is a commitment to deliver relevant information that will promote and sustain economic and social growth.

These data, information, and knowledge so fundamental to the information economy offer additional value and greater applicability when they can be represented spatially. Knowing, for example, how vector and water borne diseases may spread or attenuate over epidemiological time is important knowledge in the public health professional's toolkit but additionally understanding how diseases can spread or contract over space—as a function of local environmental factors—yields considerably more information with which to control or otherwise abate a public health emergency. This is but one example. While the applications are many and diverse, oft-cited sectors of the economy that benefit directly and significantly from geospatial information include travel and tourism, communications, public utilities, transportation, national defense, agriculture, emergency services and public security, public health, environmental management, land administration, and resource extraction such as petroleum, minerals, and forestry.

Geoinformation is, thus, now firmly recognized as a necessary ingredient in local and national development planning, decision-making, and in the monitoring and tracking of social and economic development indicators, such as the Millennium Development Goals (MDGs). Given, then, the exigency of geoinformation products in the social and economic development value chain, it follows that all African countries should have a geographically comprehensive and high-quality repository of core geospatial data sets to inform the social and economic development process, as described above. What geospatial data sets, though, comprise such a core?

This paper presents the main findings of a user needs survey undertaken across Africa for the purpose of identifying a continent-wide common and consistent set of key geospatial data. The survey was in response to a tender issued by the Chief Directorate of Surveys and Mapping of the Government of the Republic of South Africa.

The tender was part of the *"Mapping Africa for Africa"* (MAFA) initiative, launched by the Committee for Development Information (CODI) Subcommittee on Geo-information (Geo) of the United Nations Economic Commission for Africa (UN-ECA) in collaboration with the International Cartographic Association (ICA). MAFA aims to address the issue of a lack of accurate, reliable and up-to-date fundamental geo-spatial data sets essential for effective and efficient decision making and development planning in Africa. As part of this study it was necessary to determine what makes up fundamental geo-spatial data sets from a user perspective.

The objective of the study was to undertake a user needs analysis and to determine the following:

a) What is deemed to be the fundamental geo-spatial data sets (at national and sub-regional and regional level), from the universe of geo-spatial data sets, using criteria to be agreed upon?
b) For each fundamental geo-spatial data set, what spatial and descriptive (non-spatial) information is required to be collected and maintained, including the level of detail (spatial resolution and semantic level), accuracy and metadata?
c) Any temporal requirements to meet application needs (i.e. how up to date the data set must be, or the time intervals between the revision of the data set).

The detailed terms of reference for the study is presented in Annex 1.

The study was undertaken by a consortium led by South Africa's Human Sciences Research Council (HSRC) working in collaboration with EIS-AFRICA, a pan-African non-governmental organisation that seeks to promote the use of geo-spatial information in sustainable development in Africa. The consortium included other key institutions and individuals from across the continent involved in various aspects of geo-information applications, training, capacity building, and research.

This paper consolidates inputs from different perspectives and makes recommendations on "candidate" fundamental data sets for Africa, taking into consideration the different inputs.

2. Problem Statement

Mapping of the African continent has been at best very patchy. The territories of many countries have not been systematically mapped, particularly in the post-colonial era, at scales that are adequate for national development purposes. There have been several project-specific mapping activities, but they have often been sporadic and have usually tended to meet the minimum requirement of a particular project. In addition, institutions in the North hold much of the data, with little or no access to users in Africa.

The New Partnership for Africa's Development (NEPAD) and the Millennium Development Goals (MDGs) provide both policy; strategic- and programme-level frameworks for addressing Africa's development in a coherent manner. In order to achieve these noble goals it is necessary to develop a well-structured and comprehensive data foundation that would be consistent, comparable and compatible at the local, regional, national, and global levels. Such a foundation would identify and reconcile the common and key sets of information for development across the continent. A continent-wide initiative such as NEPAD provides the policy-level demand to address Africa's mapping needs comprehensively.

The problem of determining fundamental data sets for Africa can be broken down into a number of key issues, including the following questions:

- i. What are fundamental data sets?
- ii. What data sets are available in Africa?
- iii. How can the missing data sets be provided?
- iv. Should there be data standards?

This report concentrates mainly on the first question, that of using criteria to define what are the fundamental data. It provides a basis and focus for addressing the availability of fundamental data sets in the long term.

3. General Approach to Study

Organisations within the different sub-regional economic communities of Africa were carefully identified to assist the HSRC and EIS-AFRICA in undertaking the user needs analysis. These included the Southern Africa Development Community (SADC) Regional Remote Sensing Unit (RRSU) for Southern Africa; the Regional Centre for Training in Aerospace Surveys (RECTAS) and the AGRHYMET Regional Centre for West Africa; the Regional Centre for Mapping of Resources for Development (RCMRD) for East Africa; l'Association pour le Developpement de l'Information Environnementale (ADIE) for Central Africa and the Centre for Environment and Development for the Arab Region and Europe (CEDARE) for North Africa. For operational convenience and ease of reach, countries were allocated to the organizations based primarily on their mandates, which countries the organisations were already working with and to ensure a complete coverage of the continent (Annex 2).

In order to bring a global and regional academic perspective to the study, the Geomatics Division of the University of Cape Town, and the International Institute for Geo-information Science and Earth Observation (ITC) of The Netherlands were also requested to make inputs. These contributions provided historical, societal and theoretical perspectives essential to contextualise the evolution of needs and issues with regard to developing fundamental geo-spatial data sets for Africa as a whole and in the international community.

The study was undertaken through a desk study that was supported by a comprehensive questionnaire (in English and French) and telephonic interviews with key informants, with emphasis on the interviews. The regional and sub-regional partners identified were responsible for conducting literature reviews and preparing reports on user perspectives and needs in the different regions of the continent.

The questionnaire was structured essentially to serve as a guide for the interviews and partners were encouraged to explore key topics with the different people interviewed. The questionnaire was divided into five sections namely, personal details, institutional description, criteria for determining fundamental geo-spatial data sets, metadata required and spatial attribute features that make up fundamental data. One of the questions requested respondents to identify and rank, in the order of importance on a 1-5 scale (with a value of 1 being least important and a value of 5 being absolutely critical), data sets that they considered to be fundamental at the national, sub-regional and regional levels.

The questionnaires were sent by e-mail to the identified informants ahead of the telephonic interviews. Allowance was made for situations where it was recognised that it would not be feasible or possible to hold the entire interview telephonically. In such situations the minimum that was required of the partners was for a telephonic contact to be made with the key informant to discuss the questionnaire generally, clarifying issues, and then leaving the respondent to fill in and return the completed questionnaire. A number of institutions were identified and contacted in each sub-region, with the aim being to include at least five ministries in three countries which best represent specific policy issues of the region, and four national, sub-regional or regional institutions.

In addition, EIS-AFRICA sent out the questionnaire to recipients on its mailing list. This generated responses and very useful perspectives from a variety of individuals as well as institutional users within and outside Africa.

A number of multinational organisations indicated in Annex 3 were identified and participated in study. Of the twelve multinational organisations that were identified for the study, ten were part of the United Nations. One was an international geo-spatial initiative that focuses on the provision of environmental, natural resources and surveys and mapping information (i.e. Global Mapping Initiative) and another was a surveys and mapping organisation that has had extensive experience in the provision of geo-spatial data sets for African countries (i.e. Swede Survey). These organisations cover a broad spectrum of sectors including agriculture, development, environment, health, human settlement, meteorology and humanitarian aid.

Programmes and/or individuals within each of these organisations were sought that could discuss with some authority issues relating to fundamental geo-spatial data sets and the organisation's data needs. Extensive contributions were received through telephonic interviews and by the provision of extensive literature.

A pattern analysis of the completed questionnaires and interviews was then done. A pattern analysis attempts to identifying 'common points that appears throughout' the responses received from the different organisations. The analysis focused on specific issues relating to the criteria and definitions of fundamental data sets. A summary table of fundamental data sets identified was generated from the responses received.

A point system was used to identify the key fundamental data sets from all the geo-spatial data sets listed by the multinational agencies. A value of 5 was allocated to the data set identified as the most important, a value of 4 to the next most important and so on, with the data set listed as the 5th, 6th or 7th most important being given a value of 1. The values allocated to each data set were then be added up and divided by the number of times that they were listed. The data sets with the highest values were the most important fundamental data sets identified by the multinational agencies.

From all the responses received each of the data sets were described in terms of their themes, spatial features, attributes and level of detail. Finally, an assessment was done of the data sets as to whether they matched the criteria and definitions of fundamental data sets provided by the different respondents.

4. Findings

4.1 An Academic Perspective[1]

4.1.1 Review of definitions

'Fundamental' refers to the foundation on which something is built or from which something is derived. A process, phenomenon or, as relevant here, a set of data, can be considered 'fundamental, if it is primary in a sequence of events of a process, and essential in a sense, that without it the process cannot be completed.

A review of regional and international literature revealed that there is no universally accepted or unique definition of 'fundamental spatial data'. Attempts have been made to find, and agree on, definitions for 'Fundamental Data' or similarly 'Global -', 'National -', 'Framework -', 'Base -', 'Reference -', and 'Core Data'. It would seem that all of these definitions are so generic that they do not unequivocally identify the data types belonging to these categories. Any of these data categories are 'fundamental' in relation to some subsequent process, such as the addition of themes or attributes. However, these processes are so wide-ranging and varied that they do not uniquely identify the 'fundamental data' required for their execution.

4.1.2 Data Categories

In order to arrive at a definition that would be unique, an attempt has been made to pragmatically classify geo-spatial data in a hierarchical order based on their dependence on each other and the sequence of their production. A model is also suggested for differentiating between 'fundamental' and 'non-fundamental' spatial data, and suggests a criteria or scheme for classifying fundamental data. The classification distinguishes between 'primary' and 'secondary' data.

'Primary Data' is defined in this context as data that can be derived without analysis or interpretation, other than the survey calculations required. 'Secondary Data' is thematic data that is derived from the analysis of primary data or through a process of data collection in the field, statistical data collection and/or image interpretation.

1 Extracted from the paper contributed by Heinz Ruther, University of Cape Town, Geomatics Division

They can be qualitative (e.g. areas with different farming activities) or quantitative (e.g. population counts).

Primary data can be subdivided into three levels (0, I and II), and some of the secondary data can be categorised as being fundamental (Level III) or non-fundamental (Level IV; see Table 1). The principal criterion for categorising data as fundamental in Table 1 is the interpretation of the concept of 'fundamental' as formulated above. The cut-off point for fundamental data sets, level III as opposed to IV, is chosen on the basis of data volume. This was done in the interest of keeping the fundamental data manageable. With this in mind the majority of thematic data are categorised as non-fundamental due to the large number of thematic data sets and the often-high resolution at which the data is collected. In this study the focus has solely been on data categorised as fundamental.

Table 1. Structure of Geo-spatial Data

Level 0	Survey data essential for all subsequent data sets and first in the production process. This category includes the 'base maps' for Geographic Information Systems. The processing and analysis for Level 0 data is generally restricted to geodetic calculations. Geo-spatial data in Level 0 have the highest degree of objectivity, as no interpretation is involved in their production.
Level I	Geo-spatial data, which rely on Level 0 data for their creation. There is limited interpretation (e.g., classification of a water body as 'river' or 'stream' in a topographic map), but the degree of objectivity remains high. This category also includes boundaries which are the result of a human decision processes (e.g. nature reserves) as opposed to directly manmade features such as roads and other infrastructure. Level I data are generally without attributes (other than geographic names) and manmade features.
Level II	Geo-spatial data related to manmade features. The definition of Level II data is identical to that of Level I, but relates to data on manmade features only.
Level III	Generic thematic data based on primary data and derived by analysis. Data in this category are of a thematic nature, but of general interest and essential for other thematic data.
Level IV	Specific thematic data derived by analysis. This category encompasses all geo-spatial data not falling into Level 0 to III. The data can be qualitative or quantitative, as long as they can be spatially referenced.

The various data levels may have supporting data, required to generate the data sets at this respective level. These supporting data do not belong in the category of fundamental data but should be referenced as source data in a metadatabase. As an example, aerial photographs fall into this category unless they have been ortho-rectified, which would turn them into fundamental data.

4.2 Review of previous assessments and approaches[1]

Examples abound of the considerable effort that has been made to develop *harmonised and internally consistent data sets* for various applications in Africa. However, these efforts stopped short of defining a universal set of data that would be useful as a basis for all applications across the continent.

4.2.1 User Needs Assessments in Africa

Africa saw an impetus and proliferation in geo-information production and management projects in the mid-1970s in the wake of the earth resource satellite programmes. Many of these initiatives sought to improve availability of natural resource information and its use in decision making processes. A few of these have been very successful, but most of them remained project-focused, and did not address long-term integrated development information needs.

The shift in thinking regarding the environment and sustainable development towards the end of the 1980s brought in its wake a new demand for environmental information. In 1990, a World Bank Technical Paper (Falloux, 1990) formed the basis for the launch of the first broad-based multi-donor effort, supported by a broad coalition of stakeholders, to improve the availability of land-related information across Africa. It led to the launch of the Program on Environment Information Systems in Sub-Saharan Africa (The EIS Program), the first Africa-wide initiative to facilitate capacity building in spatial information management. This was developed in the context of the National Environmental Action Plan (NEAP) or National Conservation Strategy (NCS) processes, at a scale sufficient to have a long-lasting impact (Gavin and Gyamfi-Aidoo, 2001).

The EIS Program pioneered a *demand-driven approach* requiring that the production of information had to correspond to priority needs of users at various levels. In order to achieve this, *the assessment of need* had to start with an understanding of the decisions to be made, the context within which such decisions would be made, and the level at which the decision maker functions. On the other hand, users had to be able to articulate their needs clearly. Against this background user needs assessments for "environmental information" was undertaken in most African countries as part of the NEAP and NCS processes that provided the primary context in the 1990s for developing essential geo-spatial data sets to support environmental management. Need assessments were undertaken in Burkina Faso, Côte d'Ivoire, Eritrea, The Gambia, Ghana, Nigeria, Uganda, Senegal, and Tanzania, to mention a few well-documented ones.

Efforts by neighbouring countries to collaborate spawned off several sub-regional initiatives for which assessments at the sub-regional level were undertaken. Notable among these were assessments covering the entire Southern Africa Development Community (SADC) in 1993 under what became known as the SADC-EIS Programme; Central Africa in 1998 under the Regional Environmental Information Management Programme; and

1 Extracted from the paper contributed by Jacob Gyamfi-Aidoo and Sives Govender, EIS-AFRICA

the Regional Integrated Information System under the Inter-Governmental Authority on Development in 1999. More recently, in 2001, another user needs assessment was undertaken in the SADC sub-region for the Program for Regional Information Sharing and Management on Environment and Sustainable Development.

4.2.2 Core environmental data sets

The EIS initiative spun off many new *mandate-related* initiatives. For instance, in recognition of the growing need for core data sets to support regular comprehensive environmental assessments and reporting, the United Nations Environment Programme (UNEP) recommended a general guidance on what could be considered *as core data for environmental assessment purposes at national and international scale,* including the following geo-spatial data sets: *land use/land cover, hydrology, infrastructure, climatology, topography and soils* (UNEP, 1994).

On the basis of this several national user needs assessments were undertaken by UNEP under the *Environment and Natural Resources Information Network* programme, to catalyse and assist in national-scale capacity building in environmental assessment and reporting.

The United Nations Institute for Training and Research also undertook a series of user needs assessments as part of the implementation of the *Environmental Information System on the Internet* initiative to facilitate the integrated management of data and information to implement multilateral environmental treaties on desertification, biodiversity, climate change, and wetlands.

4.2.3 The Africa Data Sampler

Another spin-off of the EIS Program was the growth by the mid-1990s in demand for, and an increase in the capacity to use digital spatial data in several African countries. However, the paucity of mapped information, especially up-to-date national base and thematic maps, was a major stumbling block. In response to this need, the World Resources Institute developed the *African Data Sampler* (ADS) as a prototype database in 1994 to provide a set of *internationally comparable maps* at the scale of 1:1,000,000 for all 53 African countries (World Resources Institute, 1995).

The development of the ADS is mentioned here because it was perhaps the first attempt to provide an *integrated, comparable, and consistent set of geo-spatial data* for the whole of Africa. The objective of this *integrated spatial database* was to increase the availability of standard data, thereby providing a spatial tool for high quality presentation and reporting in a decision-making context. It presented consistent data sets for each country, and covered the following themes: *major road and rail networks, hydrologic drainage systems, utility networks (cross-country pipelines and communication lines), major airports, elevation contours, coastlines, international boundaries, and populated places.*

The ADS was based on the Digital Chart of the World (DCW) whose primary source of data was the 1:1,000,000-scale Operational Navigation Chart (ONC) series. Data sets representing the various ONC themes were "clipped" for each country. The ADS database included data on *protected areas, forests, mangroves, wetlands and sub-national administrative boundaries* with corresponding population estimates.

4.2.4 AFRICOVER

In response to a growing demand by African countries for reliable and *geo-referenced information on natural resources* at sub-national, national and regional levels the Food and Agriculture Organisation (FAO) also launched an initiative called AFRICOVER in 1994. Analysis of national needs had indicated a need for reliable and *homogeneous basic geographic information*, showing both the usual landmarks and land cover. The purpose of AFRICOVER was therefore to produce the basic geographic information common to the information components of actual and future programmes on natural resources in African countries (FAO, 1998).

The AFRICOVER initiative was therefore designed to establish for the whole of Africa, a digital geo-referenced database on *land cover and a geographic referential base* including: *geodesy, toponomy, roads, and hydrography*. The initiative was launched in East Africa covering ten countries and was implemented during the period 1995-2002. A Multipurpose AFRICOVER Database for the Environmental Resources (MADE) has been produced. In addition the project has developed an innovative land cover classification methodology, which has now been adopted by FAO and UNEP as the standard land cover classification system for the Global Land Cover Network (GLCN). Databases for each of the countries have been completed (FAO, 2005).

4.2.5 Examples from SDI initiatives

While applications of geo-spatial data vary, practice shows that most users have a recurring need for a few types of data. A comparison of fundamental data sets for various countries and programmes discussed below is summarised in Table 2.

Mexico[1]

The Mexican Spatial Data Infrastructure (SDI) adopts a definition and characterises fundamental data as follows:

- *data for which there is a basic necessity*, having different degrees of coverage (local, national, regional, global);
- sets of geo-spatial data that constitute the *foundation* for the production of added value information, applications development and the acquisition of other data.

1 This section is condensed from two papers by Francisco A. Hansen Albites, Geodesy as a Fundamental Data Set in the Mexican SDI (IDEMEX), From Pharaohs to Geoinformatics, presented at the FIG Working Week 2005 and GSDI-8, Cairo, Egypt April 16-21, 2005, and a paper entitled A characterisation of data in the context of SDIs (http://gsdidocs.org/gsdiconf/GSDI-7/papers/TStgFH.pdf, accessed 13 June 2005), originally prepared by the same author for GSDI-7.

Fundamental data is that data that is the core or the common denominator of all geo-spatial information sets, as well as the minimum required to spatially represent a given theme. In other words, fundamental data are those *data sets without which it is impossible to construct logical, consistent, accurate, rational and interchangeable geographic information.*

Table 2: Comparison of fundamental data sets for various countries and programmes

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Transportation[1]	◆	◆	◆	◆	◆		◆	◆	◆	◆	◆	◆
Administrative boundaries	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
Hydrography	◆	◆	◆	◆		◆	◆	◆		◆	◆	◆
Settlements/Population Centres		◆	◆								◆	
Topography/Physiography	◆		◆			◆	◆		◆		◆	
Elevation/Hypsography	◆	◆	◆	◆		◆	◆			◆	◆	◆
Vegetation	◆	◆	◆									
Land Cover		◆			◆		◆	◆		◆		◆
Land Use		◆					◆	◆			◆	◆
Geodetic Control	◆			◆	◆	◆	◆		◆		◆	◆
Cadastre and Tenure	◆			◆	◆	◆	◆		◆	◆	◆	◆
Imagery[3]	◆			◆		◆	◆		◆	◆	◆	◆
Cultural Environment	◆				◆							
Socioeconomic Data					◆							
Geographic/Place Names						◆			◆			◆
Geology							◆	◆				
Demography							◆					
Coastlines			◆			◆			◆			
Property Street Address									◆			
Freehold and Crown Tenure									◆			
Electoral Boundaries									◆			
Baselines, Territorial Sea-lanes									◆			
Utility networks			◆							◆		
Bathymetry			◆									

1 Includes major road networks, road centreline, rail networks, and airports
2 These "data sets" are identified under 5 broad "feature" classes.
3 Includes all types of imagery (i.e., aerial photography, digital orthophoto images, and satellite images

In Mexico seven groups of data were identified as constituting the fundamental data set, comprising geodetic references; *aerial photography and satellite imagery; data about relief, including DEMs; hydrographic network; communications and planimetric features*; international, state and municipal, including, coastal boundaries; cadastral data; geographic names data.

Australia and New Zealand[1]

The Department of Land Information of the Government of Western Australia defines a fundamental data set as one that *cannot be derived from another data set* and is essential to the outcomes of a number of agencies. According to the Australia-New Zealand Land Information Council (ANZLIC) fundamental data sets are those *which are collected as primary data sources, and from which other information is derived by integration or value-adding* (ANZLIC, 1996). More than one government agency requires consistent national coverage of such data in order to achieve their objectives, and it must conform to a set of standards that ensures that it can be combined with other components of the NSDI to create value-added products.

Nigeria[2]

The draft Nigerian Geo-spatial Data Infrastructure (NGDI) Policy defines a fundamental data set as *a data set with national coverage needed consistently by more than one government agency in order to achieve their objectives, that cannot be derived from another data set* and other agencies derive significant benefit from using it. A *variable number of data layers* may be considered to be of common-use and of national or trans-national importance and referred to as "fundamental". Thus, the list of the fundamental data sets should be seen as *dynamic* so that a data set that was not initially considered fundamental may later become desirable for inclusion in the list. Conversely, a data set that was initially included may later be dropped.

The NGDI identifies the following as fundamental data sets: *geodetic control database; topographic database/DEM (at the scale of 1:50,000 pending availability of 1:25,000 national coverage); digital imagery and image maps; administrative boundaries; cadastral databases; transportation (roads, inland water ways, railways, etc.) data; hydrographic data; land use/land cover data; geological database and demographic database*. These may be revised periodically as national needs change.

1 This section is compiled from the following sources (web pages): Department of Land Information, Government of Western Australia, http://www.dli.wa.gov.au/corporate.nsf/web/Fundamental+Data sets?OpenDocument, accessed 13 June 2005; Australian SDI: Evaluation of the Local Government Information Framework in a Multi-State Environment; http://www.icsm.gov.au/icsm/asdi/index.html; accessed on 13 June 2005.

2 Draft Nigerian Geo-spatial Data Infrastructure Policy, Federal Ministry of Science and Technology, Abuja, September 2003.

Namibia[1]

Namibia adopted a definition and identified key fundamental data sets similar to Nigeria. However, the data sets are expanded to include *data about natural resources and the environment, administrative boundaries, and population distribution at the national level.* In addition, Namibia's definition of fundamental data is contextualised for "the further development of the infrastructure of the country as well as for the realisation of economic, social and environmental benefits".

Namibia's draft Spatial Data Sharing Policy also specifies the scale for fundamental data and stipulates that they are to be: *"captured at a scale enabling the user to work with those data sets at a scale of 1:250,000"*. The policy also makes provision for revisions "in accordance with future national needs".

Nepal[2]

The National Topographic Data Base (NTDB) constitutes the fundamental data sets of Nepal. It contains different layers such as *geodetic data, administrative boundaries, transportation networks, buildings, hydrography, topography, utilities, land cover, toponomy and designated areas* organised at sheet level. The basis of the NTDB is the digitisation of *topographic base maps* at a scale of 1:25,000 for the Terai (Plain Areas) and the middle mountains at a scale of 1:50,000 for high mountains and the Himalayas of Nepal. A large-scale 1:5,000 to 1: 10,000 *ortho-photo database* is also provided for densely populated urban and semi-urban areas.

South Africa[3]

The South African Spatial Data Infra Structure Bill (2003) defines *base datasets* as 'those themes of information which have been captured or collected by data custodians'. As with other definitions found in the literature, this one embraces practically all data associated with coordinates or a geographic location, provided they are collected by a 'custodian' as defined in the bill. This definition is extremely wide ranging and does not discriminate between different levels of data relevance.

The National Spatial Information Framework (NSIF) adopts "framework" data sets, defining them as those *themes of geographic data that are produced and used by a large proportion of organisations and have widespread usefulness*. The NSIF has identified several geographic data themes as representative of South Africa's framework data through a series of workshops with the geo-spatial community. Seven geographic data themes have been identified through this process as indicated in Table 2. Implicitly the NSIF recognises the dynamic nature of framework data, and states that they "will continually evolve and improve."

1 Namibia's Spatial Data Infrastructure - Draft Spatial Data Sharing Policy, July, 2003

2 Rabin K. Sharma and Babu Ram Acharya, Spatial Information Management Promoting Sustainable Development, presented at the 3rd FIG Regional Conference, Jakarta, Indonesia, October 3-7, 2004, p.3)

3 NSIF web site, www.nsif.org.za

Other countries and initiatives

The SDI Cookbook[1] suggests the use of the term *framework information*, for they provide a framework of base, common-use geo-spatial information onto which other information can be portrayed. The framework represents a foundation on which user groups can build by adding their own detail and compiling other data sets.

The National Spatial Data Infrastructure (NSDI) of Indonesia [2]identifies *fundamental data sets as comprising of the geodetic framework; topographic databases, cadastral databases and bathymetric databases.* For the Ukraine, fundamental data sets mean the *geospatial information that is widely used and useful for the country.* The Ukrainian National Geo-spatial Data Infrastructure adopts what is generally accepted by "the majority of specialists in the world" as fundamental data sets, comprising nine data sets indicated in Table 2[3].

Various examples of data sets are described for The Netherlands as constituting fundamental data sets[4]. These are nationwide data sets including a *large-scale base map* of The Netherlands and a 1:10,000 core (topographic) database covering the whole country, *land cover* database, land cover *ecological* database, *waterways, geology, archaeology, cadastral map, and a digital elevation model.*

The Global Map is a project with the objective to develop digital geographic information at a scale of 1:1,000,000 (i.e., in 1 km resolution) covering the whole land with standardised specifications and available to everyone at marginal cost. Nearly 150 countries are now supporting the Global Map project. It contains eight layers based on data sets that were built as part of earlier initiatives on a global scale such as the Global 30 Arc Second Elevation Data Set (GTOPO30), the Global Land Cover Characteristics Database, and VMAP Level 0.

For the United States NSDI the *framework* includes seven groups, and Colombia lists seven *fundamental* data layers (Hansen Albites, 2004; see Table 2).

The VMAP series

The Vector Map (VMAP) is designed to provide a consistent global coverage of vector-based geo-spatial data at low (Level 0, VMAP0), medium (Level 1, VMAP1), and high (Level 2, VMAP2) resolution. The content of the VMAP0 database is an updated and improved version of the DCW, augmented with low-resolution bathymetry for global coverage, thus providing vector-based geo-spatial data that can be viewed at a scale of 1:1,000,000.

1 GSDI Cookbook, Version 2.0 25 January 2004, page 17

2 http://www.bakosurtanal.go.id/

3 http://www.geomatica.kiev.ua/project/nsdi/basemap_e.shtml

4 Spatial Data Infrastructures in The Netherlands: State of play Spring 2003 — Country report on SDI elaborated in the context of a study commissioned by the EC (EUROSTAT & DGENV) in the framework of the INSPIRE initiative, August 2003; Spatial Applications Division, K.U.Leuven Research & Development, Vital Decosterstraat 102, B-3000 LEUVEN, http://www.sadl.kuleuven.ac.be

The vector database is organised into ten thematic layers (i.e., the DCW themes together with bathymetry). VMAP Level 0 includes an index of geographic names to aid in locating areas of interest[1]. VMAP1 data correspond to the geometry and contents of maps at a scale of 1:250.000, and the VMAP2 database contains information roughly equivalent of maps at a scale of 1:50.000. The VMAP1 and VMAP2 databases consist of natural and man-made point, line and area objects subdivided into the following ten themes: *boundaries, data quality, elevation, hydrography, industry, physiography, population, transportation, utilities and vegetation*. The VMAP family is completed by the higher resolution *Urban VMAP* data[2].

4.3 Global perspective[3]

4.3.1 Guiding principles

Geo-spatial data sets can only be considered as fundamental if they fulfil certain conditions. The United States Federal Geographic Data Committee (FGDC) has provided the following guiding principles for building fundamental geo-spatial data sets:

- The data should be a preferred data source.
- It should represent the best available data for an area - the most current, complete, and accurate data.
- The data should be widely used and useful. Users must be able to easily integrate the framework data with their own and provide feedback and corrections to framework data.
- Access to framework data should be at the lowest possible cost and without restrictions on use and dissemination. The framework is a public resource.
- Duplication of effort should be minimised. Sharing the development and maintenance of framework data reduces the costs of individual users' data production.
- The framework should be based on cooperation. It is built through the combined efforts of many participants who work together on its design and development and contribute data to it.

4.3.2 European initiatives

There are two major initiatives on fundamental geo-spatial data in Europe. These are EuroGeographics and INSPIRE. EuroGeographics is building the European Spatial Data Infrastructure, with the vision to achieve interoperability of European mapping and other geo-spatial data. The INSPIRE initiative aims at the creation of a European spatial information infrastructure to provide integrated spatial information services.

---

1 See http://store.geocomm.com/viewproduct.phtml?catid=25&productid=1194, viewed 21 May 2005.

2 See Ohlhof, T., et al., Generation And Update Of Vmap Data Using Satellite And Airborne Imagery; http://www.ipi.uni-hannover.de/html/publikationen/2000/heipke/1305_paper.pdf; also http://www.google.com/search?q=cache:c8_RgTYyFPYJ:www.ipi.uni-hannover.de/html/publikationen/2000/heipke/1305_paper.pdf+VMAP&hl=en&start=37, viewed 21 May 2005)

3 Extracted from paper contributed by Drs. R.A. Knippers, et al., ITC

The Working group on Reference Data and Metadata of INSPIRE, the European spatial information infrastructure, has agreed on the following reference data layers for the European spatial data infrastructure:

- Geodetic reference data
- Units of administration
- Units of property rights (parcels, buildings)
- Addresses
- Selected topographic themes (hydrography, settlements, transport, height)
- Ortho-imagery
- Geographical names

The European Territorial Management Information Infrastructure (ETeMII) project proposes similar layers of information as fundamental geo-spatial data and EuroSpec have used these data layers for investigating the possibility to make specifications for a European reference data set.

4.4 Perspective from multi-national agencies[1]

4.4.1 Criteria for defining fundamental geo-spatial data sets

There are differences in the response from the multinationals on what the criteria should be in defining fundamental geo-spatial data sets. However, there seems to be agreement that the *foundational* aspect is the most important factor. "Foundational" in this context means those *geographically referenced features or objects* of an area that are generally found on topographic maps and used as a base to build other thematic and core data sets. These features or objects are defined by their spatial dimension and the distinct attributes associated with them. They tend to be fairly stable over time but fundamental data sets should also ensure that changes to geographic features or objects over time are reflected.

A key aspect of being foundational is that they provide the *geographic reference* base upon which other layers of geo-information can be developed. A term commonly used to describe fundamental data sets is *primary*. In this context "primary" is meant as those fundamental data sets that are the first component of a process needed to develop other data sets or applications. They also form the base on which other layers of geo-information can be overlaid or integrated to possibly produce new layers of information.

Key characteristics of fundamental data sets are their *standards* that should be uniform*ly* applied across the world. This is required to enable the fundamental data sets to be used within and between countries. The FAO is of the opinion that *consistency, quality, continuity, and accuracy* are key characteristics of fundamental data sets. The World Health Organisation (WHO) adds that high levels of validation should characterise fundamental data sets.

1 Extracted from paper contributed by Craig A. Schwabe, HSRC.

4.4.2 Definitions of fundamental data sets

Although the definitions of fundamental data sets provided by the multinational agencies were diverse they also reflect similar themes. Common themes that could be identified in the definitions provided included application and use, coordination, referential base, spatial data and standards. The *application* aspects of the definitions emphasised the use of fundamental data sets in many different sectors by a diversity of users for decision-making purposes. Although not explicitly stated, it is implied that fundamental data sets should be used within an appropriate *framework*.

Another key theme highlighted in the definitions was that of fundamental data sets being a *spatial reference base*. The definitions emphasise that fundamental data sets are the basic layers of geo-referenced data that contain information on features, objects, elements and/or entities that are located on the surface of the earth. As the Global Mapping Initiative puts it they are basically the geo-information found in topographic data sets. What is again implied in the definitions is that fundamental data sets are those upon which other thematic and core data sets are developed.

The UNECA states that fundamental data sets are those that are identified by a 'responsible coordinating body or accepted by the user community'. This statement emphasises the necessity for *coordination* in the development of fundamental data sets and *consultation* with a broad audience of users.

4.4.3 Data rankings

Using the scoring method described in section 3, various data sets were scored and ranked as shown in Table 3.

Table 3: Summary table of fundamental data sets

Data Set	National	Sub-regional	Regional
Administrative boundaries	16	20	18
Geodetic reference	15	10	15
Transportation	12	15	12
Human settlements	10	10	5
Hydrography (i.e. rivers, lakes)	7	9	12
Geographic place names (e.g., cities, towns etc.)	7	7	7
Elevation	7	9	9
Land cover	6	7	8
Temperature	5	5	5
Rainfall	4	5	7
Population	4	5	5
Land use	4	4	4
Health facilities	4	4	0
GPS base stations	4	0	0
Vegetation cover	3	2	2
Soil type	3	1	1
Relative humidity	3	3	3
Infrastructure	3	0	0
Cadastral information	3	3	3
Solar radiation	2	2	2
Wind speed	1	1	1
Imagery (e.g., ortho-photos, satellite)	1	1	5
Health districts	1	1	0

4.5 Sub-regional perspectives

4.5.1 East Africa[1]

In the East Africa region, the definition and development of common fundamental data sets is at the infancy stage, and not much research has been conducted into the issues of fundamental data sets with an East Africa perspective. For the purpose of this study major users or players in the development of geo-spatial data in Kenya, Uganda, Tanzania Ethiopia, Somalia, Sudan, Mauritius and Djibouti were targeted. The major players were the government ministries of environment, lands and transport, academic institutions and private mapping organisations. In addition to this group, participants from the region taking part in a 3-week training course on use of geo-information for environmental assessment and reporting were also interviewed for the study.

Responses were obtained from six of the eight countries surveyed: Kenya, Tanzania, Uganda, Ethiopia, Mauritius and Somalia. The common points identified by respondent

1 Extracted from paper contributed by the Regional Centre for Mapping of Resources for Development

from all the sectors with regards to the criteria for defining fundamental data sets are *coverage, importance to a wide variety of users, availability/accessibility and reliability or accuracy and scale*. In other words, respondents considered fundamental data sets to be those that *consistently cover the whole country and are readily available to a wide range of users*. Another important criterion identified was the use of such data for *orientation*.

The most common definition of fundamental data sets given by respondents is that they are *basic data that can be used as a reference base by a diversity of users in many application areas*. Also fundamental data sets are a *base on which other thematic data may be linked to the landscape*. Data sets listed as fundamental were scored and ranked, with those with the highest values being considered the most important. The following data sets were identified, in order of importance:

- Topography
- Landuse/cover
- Administrative boundaries
- Drainage
- Transportation

Each of these data sets were described in terms of their themes, spatial features, attributes and level of detail in the report received from Eastern Africa.

4.5.2 West Africa[1]

Questionnaires were sent to 38 potential respondents in 14 countries (Benin, Burkina Faso, Cameroon, Chad, Côte d'Ivoire, The Gambia, Ghana, Guinea, Liberia, Mali, Niger, Nigeria, Senegal, and Sierra Leone). However, after concerted efforts through telephone and email contacts, responses were received from only 6 countries — Benin, Burkina Faso, Ghana, Liberia, Nigeria and Senegal. The response rate was about 32 per cent, with a third of the respondent organisations being government ministry/department. The rest were local authorities, non-governmental organisations, semi-governmental and parastatals, and academic/research institutions. Respondent institutions had been selected on the basis that they best represented geo-information activities in the respective countries.

The respondents defined fundamental data sets from different perspectives. Prominent among the definitions were: *"dataset with national coverage needed consistently by more than one user"; "that which gives a thematic data developer the basis to develop his own"; "a variable number of core data set that is of common use having national or trans-national importance needed consistently by more than one government agency in order to achieve thematic data"*.

In the context of MAFA initiative, respondents believe that fundamental data sets should have the following characteristics:

1 Excerpted from paper contributed by the Regional Centre for Training in Aerospace Surveys (RECTAS).

- required by more than one organisation (GI users or providers);
- impossible or difficult to derive from another dataset;
- many users or providers derive substantial benefits using the dataset;
- necessary for many applications; and
- collected by government agencies.

For the purpose of the analysis, a data set was considered fundamental for each country when 50 per cent or more of the respondents selected the data set in the respective country.

A summary of the responses is: administrative boundary, populated places/settlements, topography, hydrography and transportation. The following metadata attributes were selected for the fundamental data sets: *originator of the data set, contact details of custodian, date of publication, accuracy of attributes, title, accuracy of spatial data, format of the data set, scale of maps, description of the data set, projection/coordinate system, purpose of collection, datum, ellipsoid, date of completion, access constraints, completeness, use constraints and spatial boundary extent.*

The order of importance of datasets at the sub-regional level is summarised below:

[illegible]	[illegible]
High	Administrative boundary
	Populated places/settlements
	Topography
	Hydrography
	Transportation
Medium	Populated places/settlements
	Agriculture/Forestry
	Vegetation
	Geodetic controls
Low	Cadastral
	Land use
	Population data
	Minerals
	Poverty, Health, Security

The relevance of fundamental data sets in terms of social and economic development did not come out from most of the responses. However, it is worthy of note that some respondents see socio-economic data (presented geographically, e.g., health, security and education) as fundamental data sets, if the meaning of fundamental is "the most important or essential data sets". These data sets are of course important and basic to the growth and development of any society, but the presentation of such data sets is based on another data set.

In the context of poverty alleviation in Africa it is important to consider cadastral and property related data sets as fundamental dataset. It has been argued that landed property whose economic and social aspects are not recorded in a formal property system is really hard to move in the market to promote investments (de Soto, 2000).

Therefore, though it does not emerge from the survey and the threshold used in the study that cadastral and property related data sets as fundamental, we argue that, from the perspective of poverty alleviation and economic empowerment, these data sets are in fact the data that everybody needs and the foundation for all other data sets. Large-scale cadastral maps may be the bases for other topographic information.

In addition to the "conventional" definition, fundamental data sets may also be viewed from the point of view of national concern and the data sets that every body needs. We are of the strong opinion that fundamental data sets should be determined at national level as a policy issue on a dynamic basis. If fundamental data sets were basic and essential data sets for social and economic development, then they would obviously change with the social and economic requirements of the country. Thus, once determined there should be a continuous exercise to review of a country's fundamental data sets.

4.5.3 North Africa[1]

The data for the study was gathered using a questionnaire provided by the HSRC, and which was designed primarily for primary data, i.e., first hand information gathered from original sources. The questionnaire was mailed to the participating institutes. Researchers of CEDARE used their own judgement about which respondents to choose, and picked only those who best meet the purpose of the study.

Institutions from Egypt, Sudan, Tunisia and Morocco were selected to participate in the study. Five of these were governmental institutions (ministries and/or public departments). The others comprised a local authority, a semi-governmental/para-statal organisation, an academic and research institute, a private company, an international organisation, and an inter-governmental institute organisation. The respondent institutions operate at different levels (geographic coverage — from local administration to international), and functional application areas (sectors). Thus, the observations and conclusions from the study are sufficiently heterogeneous and representative of the community that the study addresses.

Naturally, responses varied depending on the mandate, geographic scope, the functional application area focus, clients, and scope of work of the respondent institutions. There was a wide spectrum of perspectives as to what constitutes fundamental data. In essence, each respondent organisation presented its own adopted terminology and definitions to suit its needs. While some focused on how critical some geospatial data is to their "classical" users, others have focused on the inter-operability issues. Such diversity

1 Excerpted from paper contributed by the Centre for Environment and Development for the Arab Region and Europe (CEDARE)

may reflect the challenges each of the respondent organisations face when attempting to harmonise their datasets (and possibly operations) with others.

The report highlights a dichotomy regarding cartography and GIS in the region. There has been a trend towards a GIS-centred view of fundamental datasets. In this context the inter-operability of fundamental data has become a major issue and as a result is mentioned repeatedly, in various ways, under different conditions and geographic extent. This reflects a change in the vision of some in the region, and the two "views" (classical cartography versus GIS) shaped some of the responses in terms of ambitions and more honest views of actual internal processes. Suggested definitions of fundamental datasets clearly reflect these differences. Respondents' priorities of datasets can be prioritised as follows:

[illegible]	[illegible]	[illegible]	[illegible]
First	Topographic	Irrigated/non-irrigated lands	Transportation
	Water ways	Water quality status	Water
Second	Geology	Topographic	Topographic
	Soil	Geology	Geology
	DEM	Soil	Weathering
	Land use and cover	Environmental	Thematic
	Environmental	Weathering	DEM
		Land cover and use	
Third	Cadastral		Soil
	Weather		Environmental data
	Coastal zone management		
	Socio-economic		
	Derivative		

The differences extend to suggested datasets, their classes, and themes. They also extend to level of details, planning purposes, and how often the data set should be updated. However, the results suggest there is an association between the level of detail and usage. Responses regarding the criteria and definition for fundamental data sets vary according to the level. For instance, the data set on topography is considered absolutely critical at the national level, but importance declines once the use of the information is at sub-regional and regional levels; it becomes "important" at these levels. Similarly, data sets on roads, property, road intersection, bridges and grazing lands seem to be "very important" at the national level, as are data sets for geology, soil, environment, waterways and water, and land use. However, they are considered just "important" at the sub-regional and regional levels. Data sets that seemed absolutely critical and very important at the sub-regional and regional levels include simple land use, political boundaries, rivers and lakes, irrigated and non-irrigated lands, status of water quality. Respondents indicated that wetlands, desert lands and watershed area are very important at the sub-regional level.

It needs to be pointed out that answers to some of the questions in the questionnaire suggest a problem with semantics. Clearly this arises from the challenge of defining data sets and their significance, not only across languages, but also across geographic zones. Even when geographic entities in one language are 'translated' to the nearest 'label' in another language the resultant "entity" might be fundamentally different as natural entity. Even man-made features pose the same problem, even among organisations of different jurisdictions of the same country, such as in the case of United States. Therefore, a "Don't Know" response might actually indicate a 'hesitation' by the respondent to associate a data set to a 'similar' but not quite the same one mentioned in questionnaire.

Another important point worth noting is that the freedom to disseminate information is relatively limited in North Africa. Be it the result of historical inertia of governmental systems, lack of resources, or even resistance to culture-specific changes, it could be said that changes are slower than expected. This is a situation that, in one way or the other, clearly affected the responses of various institutions.

The questionnaire contents also suggest a hint of a "face-lift" which might indicate that respondents were presenting themselves as possible candidates for capacity-building funds. It was clear that responses received from the institutions reflected some amount of exaggeration of roles and needs for data, and need for assistance.

Mailing the questionnaires has its own disadvantages. First, the response rate was low. CEDARE staff spent a considerable amount time and effort to retrieve the questionnaires from the participating institutions. Secondly, with no interviewer present, there was no possibility to vary the question and probe further for a specific answer if the initial response was vague. Also the absence of an interviewer meant that there was no opportunity to observe non-verbal behaviour, or to make personnel assessments. Also there was no possibility to resolve problems relating to technical semantics.

This study was obviously ambitious, and that fact itself constituted the major weakness of the study. The number of questions and details to be measured on ordinal scale makes it difficult to infer any association between the different variables, and to model this behaviour to draw solid recommendations.

4.5.4 Southern Africa[1]

Fundamental geo-spatial data was cited as one of the components in the development of a holistic national data foundation. Review of available literature showed that most SADC countries have adopted the concept and need for the implementation of SDIs in their countries. National mapping agencies of these countries have made a lot of effort to convert their data sets from analogue to digital format. South Africa, Botswana, Namibia, Swaziland and Malawi are however the only countries that have maintained any serious fundamental base mapping programmes. They have also managed to include SDI activities into mainstream government activities, thereby increasing chances of getting support for the development of fundamental data sets.

1 Extracted from paper contributed by Reuben Mavima, representative of SADC on the project team.

In a review of the development of SDI in the SADC region, Mavima and Noongo (2004) noted that the process has been hindered to different degrees and for a variety of reasons, including the lack of political support and dwindling budgets. In countries where some progress in the development and maintenance of fundamental data sets has been made, local standards have been implemented that are not compatible with international standards. It is also common that donor funded projects have resulted in data sets that became obsolete as soon as the projects end, and organisations have been stuck with data sets that are not usable. Furthermore, each country in the SADC region has had several studies relating to fundamental geo-spatial data sets. These studies have often been donor funded and project-based, and the outputs are not easily accessible to researchers.
In SADC, potential users from different sectors in the region were identified. An electronic questionnaire was distributed to thirty-four individuals from different organisations, followed by telephone calls to clarify issues that were not clear in responses to the questionnaire. The telephone calls helped to increase the response rate. The selection of contacts took into consideration those organisations that handled geo-spatial data, with a bias towards government ministries or departments. To get diverse views, regional and research institutes were also considered.

One clear message coming out of the survey in SADC is that there are diverse opinions on the definition of fundamental geo-spatial data sets. Varied responses were given for criteria that should be used to define a fundamental geo-spatial data set. The following are some of the criteria listed:

- should be baseline data that serves as a frame for other data sets
- should be basic in the sense that they are not easily derived from other data sets
- provides a common reference base for thematic mapping, (i.e. provides the context or background against which thematic maps are composed, or the background that eases the interpretation of the latter)
- should have wide applicability across all the sectors of the national economy
- should be primary/core with no dependencies on lower order data sets
- should be complete and cover the whole country
- be of permanent value

Sector applications of geo-spatial data have a bearing on the definition and ranking of fundamental data sets. However, there is a general agreement between SADC and international definitions. Using the scoring and ranking scheme described in section 3, the groups of data sets listed in Table 4 were identified, in order of importance.

The candidate fundamental data sets above were found to be in general agreement with other international determinations and are applicable to the entire SADC region. One conclusion that may be drawn from the survey is that SADC countries are generally aware of what are fundamental geo-spatial data sets. However, there are inconsistencies between countries on the naming and definition of fundamental data sets. This raises the issue of standardisation and coordination at both national and sub-regional levels.

This survey and the MAFA initiative can be used as a basis to create dialogue between countries and sub-regions on these issues.

Table 4: Data rankings for SADC

National	Sub-regional	Regional/Africa
Imagery	Geodetic network	Place names
Census	Administrative Boundaries	Geodetic network
Social Services	Land cover	Administrative Boundaries
Geodetic network	Geology	Geology
Geology	Topography	Towns
Administrative boundaries	Transport/Roads	Communication
Agriculture	Communication	Topography
Drainage	Geographic Place names	Infrastructure
Environment and climate	Infrastructure	Environment and climate
Cadastre	Environment and climate	Land cover
Topography	Population	Drainage
Land cover	Towns	Population/census
Infrastructure	Drainage	
Transport and Roads		
Population		
Towns		

Table 5 summarises and cross-references key data sets identified from the various inputs.

Table 5: Fundamental data sets identified from sub-regions of Africa

Transportation		◆	◆	◆	◆
Administrative boundaries	◆	◆	◆	◆	◆
Hydrography	◆	◆	◆	◆	◆
Settlements/Population Centres	◆	◆		◆	◆
Topography/Physiography	◆	◆	◆	◆	
Elevation/ Hypsography	◆	◆		◆	◆
Vegetation				◆	◆
Land Cover	◆	◆	◆		◆
Land Use	◆	◆	◆		◆
Geodetic Control	◆	◆	◆	◆	◆
Cadastre and Tenure		◆	◆	◆	
Imagery	◆	◆			◆
Geographic/Place Names	◆				◆
Geology	◆	◆			◆
Demography					◆
Property Street Address					◆
Utility networks	◆				◆
Climate		◆			◆
Geoid model	◆				
Conservation areas	◆				◆
Forestry reserves	◆				
Soil	◆				◆
Minerals	◆				
Ecological zones					◆
Land suitability					◆
Fauna					◆

4.6 Responses from EIS-AFRICA's Network

In order to widen the reach of the survey, EIS-AFRICA posted the questionnaire to all e-mail addresses on its mailing lists. Useful and insightful inputs were received from respondents, including a multi-national mining interest and international organisations in Europe, the USA, and Africa. The candidate fundamental geo-spatial data sets and their respective weighted scores and ranking (in order of importance) are presented in Table 6.

Table 6: Weighted scores of data sets from responses from EIS-AFRICA mailing list

[illegible]	[illegible]
Hydrological network (Drainage /rivers /water-source)	2.750
Administrative boundaries (districts, provinces, national)	2.500
Climate (rainfall, temperature, solar radiation etc)	1.875
Census (population and housing census)	1.625
Geology	1.500
Soils	1.500
Elevation (and terrain derivatives)	1.250
Land use	1.250
Population (distribution)	1.250
Road network	1.250
Agricultural census	1.125
Infrastructure	1.125
Topographic maps	1.000
Agriculture	0.875
Cadastre	0.625
Important structures/medical/school, etc	0.625
Land cover	0.625
Road/railway/Airports	0.625
Transport	0.625
Coastlines	0.625
Economic indicators	0.625
Natural resources	0.625
Ground water resources	0.625
Place names (locality, towns, country)	0.500
Socio-economic indicators	0.500
Demography	0.375
Health Demographics	0.375

5. Framework for Africa

5.1 Analysis of literature review

The review of available literature reveals that there is no universally accepted definition of, or what constitutes a *fundamental data set*. Furthermore, no universally accepted methodology is presented in the literature on how to develop such definitions or identify the fundamental data sets. Different user communities have adopted definitions and categories of data sets to suit their own needs. The literature also shows that there are differences in what various user groups identify as *fundamental*. This suggests that categories of data are identified as fundamental in response to interests in each particular instance, whether it is global, regional, national or local. Over the years different groups have adopted various terms, especially as the concept of SDI evolved, and it would seem that there is even some amount of confusion over terminology.

Data sets that may be used for many different purposes and in many different applications are variously referred to as *base* data, *fundamental* data, *foundation* data, *framework* data, *reference* data, or *core* data. Although some of these terms are used in the literature interchangeably, they may not always mean the same thing. Chapter 2 of *The SDI Cookbook* contains a discussion on the distinctions sometimes made among *core, reference, foundation and framework* data sets, and it is argued that these differences are academic (Luzet and Murakami, 2004). However, in the context of the current discussion, it is contended that these terms point to *different* aspects of geo-information which are important in attempting to define what would constitute *fundamental data sets* for an area of interest that encompasses several countries at different stages of development.

The notion of *core data* appears to relate to the *specific mandate-related applications* for which the data sets are required. Core then seems to mean *"central to"* a particular group of applications, or a set of data that is *essential for a particular purpose*, for instance, integrated environmental assessment and reporting. As is evident from the examples from various countries and programmes, it is clear that *data sets for common use* vary from one group of users to another. This was also amply demonstrated by the differences in understanding and the approach of various international organisations that were telephonically interviewed and responded to the questionnaire for this project.

5.2 Data categories

Determining fundamental data sets for the whole of Africa implies that these should be *universal* sets of data. Ideally, the data set should include *all* geo-spatial data needed to support a wide variety of applications in different contexts, at different scales from the local up to the national, regional and global levels. As a start there is need for consensus regarding a "standard" reference frame and data *integrating models* that would allow information generated by different data providers to be referenced to each other within a *coherent framework*, and which allows data to be "fitted" to each other through various processes. Fundamental data sets should comprise of the following elements:

- a ***geographic reference framework*** through which features can be located within a model of the surface of the earth;
- a reliable ***base geography***;
- a standardised ***geo-coding*** scheme for attaching geographic references to non-geographic data.

On the basis of these elements, geo-spatial data could further be classified in a hierarchical order or levels, based on their dependence on each other and the sequence of their production, in order to arrive at a *universal set of fundamental data* for Africa.

5.2.1 Geometric referencing and projection

Geo-spatial data refers to all data that can be referenced to a position in geographic space. Common to all such data is the *geometrical referencing* of features and phenomena of interest. At the base of this is the reference ellipsoid (name), with its accompanying numerical values of the ellipsoidal parameters or geodetic reference system. Information on the reference ellipsoid, height datum and map projection has become increasingly important with the expanding use of GPS equipment by non-experts.

This constitutes a *primary reference*, fundamental for integrating data from different locations and providing the physical links to a *co-ordinate system*. This reference is basic to all geographic information, and although background to, and generally not a part of the geographic information that is actually used in applications, especially Geographic Information Systems, it is of vital importance to the geo-information community. Even data from the same location but from different sources may be *re-projected* to this reference. The projection system and the underlying geodetic framework used for the production of topographic maps are inherently accepted when such maps are used as the basis for deriving other information.

5.2.2 Base geography

Base geography provides a real world, physical landscape reference, consisting of features such as the coastline. This is essential to allow the user to *relate to* or to *'refer'* external information to the real world. For instance, satellite imagery can be adjusted (rectified) using road intersections and other features identifiable on a published map. Fitting data to a reliable base geography ensures that local differences and discrepancies in geographic representation are avoided when data sets are integrated. Otherwise relationships between features may be incorrect. For instance, points that should be on the land may be in the sea, or streams that should reach the coast may not.

Other base information assists in *orientation* and reflects *spatial organisation*. Standard topographic maps integrate the geodetic framework, and usually include *base information* representing the general characteristics of the landscape (rivers, towns, hydrography, etc.) and features such as infrastructure – roads, railways, etc. Administrative boundaries and cadastral parcels would belong to this group of geo-information, providing a *basic spatial framework*, and references for integrating other information for particular locations, and also for geo referencing other data.

5.2.3 Geo-coding schemes

Spatial reference is not restricted to a discrete point, but can refer to an area or volume in space. Statistical data especially tends to be allocated to regions rather than discrete points, and *geo-coding* makes it possible to attach a geographic reference to such data. The process involves the identification of spatial locations of the data points or features using some kind of a coding scheme and relating them to their respective geographic coordinates in the chosen reference model. By this mechanism non-spatial "data points" or features can be located within the geographic reference frame, and can therefore be mapped directly.

5.3 Recommendations

5.3.1 Definition

Considering perspectives from current literature and all the varied inputs from different organisations and individuals, it is possible to propose a definition of fundamental data sets for Africa that combines key elements in the context of an initiative such as MAFA. It is recommended that the following articulation be adopted:

Fundamental data sets are the minimum primary sets of data which cannot be derived from other data sets, and which are required to spatially represent phenomena, objects, or themes important for the realisation of economic, social, and environmental benefits consistently across Africa at the local, national, sub-regional and regional levels.

The data sets must be considered as the baseline or common denominator data upon which other spatial data sets and various levels of applications are built. It is further recommended that the following guiding principles or criteria should inform the inclusion of particular data items in the fundamental data set:

- Must contain sufficient level of detail appropriate for the intended applications;
- Must include, either explicitly or implicitly, a reference frame (geodetic or coordinate);
- Must refer to or represent a place in space, or provide a context or framework for organising information in space;
- Must incorporate a clear, and unambiguous definition and scheme for representing basic information useful for common applications, including a set of key attributes;
- Must be continuous, contain consistent information, and have complete coverage for the area of interest;
- Must conform to accepted standards and norms, ensuring that it can be combined with other groups of data of any sort to create value-added products
- Must accommodate future revisions of the data set.

The definition above, set in the context of Africa's development needs, suggests that data identified fundamental data sets are dynamic and may change over time, reflecting national and regional needs.

5.3.2 Candidate data sets

On the basis of the adopted definition of fundamental data sets, the literature review, interviews, and analysis of questionnaire responses, the "candidate" data sets in Table 7 are proposed for further debate and adoption. A hierarchical order or levels of "fundamentalness" reflecting the dependence of data on each other and the sequence of their production is suggested.

Primary Reference

Geo-spatial data would be meaningless without information on the reference system and the map projection they are based on. Data on the geodetic control network is therefore the main framework for all geographic information. According to the proposed categorisation scheme the geodetic control network is primary data, at ***Level 0*** as survey data essential for all subsequent data sets and first in the production process. It provides a *common model of the earth* as the base for the data, and it is important that all data sets for common use adopt the same geometric reference in order to ensure that data can be combined and cross-referenced with one another. Developing a geodetic reference frame for Africa is therefore a *sine qua non*.

Table 7: Fundamental data sets for Africa

0	Primary Reference	Geodetic Control Network	Geodetic control points	List of coordinates with information on the history of establishment of the network as well as network design in digital map/GIS format.
			Height datum	List of heights of primary height points in digital map/GIS form (vertical datum surface)
			Geoid model	Geoid-ellipsoid separations (heights at individual points) to convert from GPS observations to heights
I	Base geography	Rectified Imagery	Aerial photography	Aerial photography
			Satellite imagery	Satellite imagery
		Hypsography	Digital elevation model	Vertical distance from the earth's surface to a base defined by the adopted height datum
			Spot heights	Heights of peaks
			Bathymetry	Vertical distance of earth's surface from base defined by Lowest Astronomical Tide
		Hydrography	Coastline	The limit of land features usually at mean high water level.
			Natural water bodies	Location of watercourses, drainage network, and all inland water bodies (streams, rivers, canals, ponds, lakes, etc.)
II	Administration and spatial organisation	Boundaries	Governmental units	Limits of administrative and jurisdictional authority (International, national, sub-national boundaries, and local government areas)
			Populated places	Population centres including urban areas, towns, localities, and rural settlements
			Enumeration areas	Boundaries of areas delineated for the purpose of collecting demographic census information
		Geographic names	Place Names	Official and local names of places
			Feature Names	Official and local names of cultural and geographic features (including roads)
		[Land management units/areas]	Land Parcels/Cadastre	A consistent framework of land parcel/cadastre boundaries defined for land tenure purposes, referenced to a common datum
			Land Tenure	Current, proposed and historical details of all tenures, e.g., details of ownership, vesting, and including traditional forms of land holding.
			Street Address	Unique Street Address of parcels/properties
			Postal or zip code zones	Boundaries of post code areas
			Land use planning zones	Boundaries of areas of permitted/restricted land use defined by planning authorities (includes conservation areas, heritage sites, and restricted areas)
	Infrastructure	Transportation	Roads	Network of physical roads and carriageways
			Road centrelines	Centreline of roads and carriageways
			Railways	Network of railway lines
			Airports and ports	Location of airports, sea ports, and navigation aids
		Structures	[Bridges and tunnels]	
		Utilities and services	Power	Locations of trunk or national grid power line networks and major assets/installations, and sources
			Telecommunications	Locations of trunk communication networks and major assets
III	Environmental Information	Natural environment	Land cover	Observed bio-physical cover over on the earth's surface1
			Soils	Boundaries and classifications of soil resources
			Geology	Boundaries and classification of geological units

1. FAO Land Cover Classification System — Classification concepts and user manual (Software version 2, Draft Version, Nov. 2004), p.7

Base geography

Base geography refers to the real world or the physical landscape. It represents ***Level I*** geo-spatial information that requires limited interpretation, and yet retains a high degree of objectivity. Imagery provides primary and basic information about the physical landscape from which a variety of geographic information may be derived. Aerial photographs traditionally served as the primary source for both large scale mapping activities and standard topographic maps. With the advent of various types of satellite imagery that are now used extensively for cartographic projects and support for GIS development, satellite data constitute a fundamental source of data. However, both types of imagery must be rectified using Level 0 data (geodetic control, ellipsoid, etc., or information that incorporates these) in order to turn them into a fundamental data set.

The physical (or natural) landscape is characterised by the hypsography and hydrography. Hypsography depicts a 3-dimensional landscape and its landforms, with the spatial features of this theme being contour lines, bathymetry lines, form lines, and spot heights. Hydrography depicts the drainage pattern, comprising the rivers, streams, canals, wells, wetlands and water bodies. It may be argued that the hydrography is naturally present and defined by the hypsography. However, the features are themselves entities on their own that need to be represented as part of the base geography.

Administration and spatial organisation

The next category of data sets appears to relate to and support the organisation and management of people, communities, society, and their activities in geographic space. These data sets arise from human decisions, or relate to man-made features. They constitute ***Level II*** fundamental data sets.

One data set in this group that occurs consistently in all the fundamental data sets that have been reviewed, as well as being indicated by all respondents, is that of administrative boundaries. Conceptually a boundary defines what may be called delimiters or the containing spaces for various categories of data items (Hansen Albites, 2004). It is clear from the various inputs that the features that define the limits of administrative and jurisdictional authority are a key data set. So too are the spaces that contain populations, either as settlements or sub-divisions delineated for the purpose of collecting information about populations and their activities.

Geographic names are particularly important. They are data elements in their own right as part of socio-cultural assets, as well as providing a means to uniquely identify features in geographic space. They are essential elements for orientation, referencing, and communication.

The rest of the data sets in this category relate to the management of land units, in terms of ownership of land as property, addressing, and use of areas or zones for specific purposes. On the basis of the adopted definition land use would be a derived data set and, therefore, is not included as fundamental. However, information depicting restriction or accessibility to a land unit as established by relevant legislation, for example, areas reserved for such purposes as conservation, heritage sites, and restricted areas, need to be published as part of the fundamental data set.

Infrastructure

This category of data relates mostly to the built environment, but focuses on transportation and service infrastructure. These are man-made features, and therefore the data sets are also categorised hierarchically as Level II.

Transportation information can be considered as part of the base geography. However, they have a specific primary human function, serving as connectors between populated places as well as functional service centres. They also provide references for integrating other information and for orientation. The transportation theme includes features such as main roads, secondary roads, minor roads, streets, tracts, etc. Railways and airports are also included under this theme.

Major utility and service networks are also represented under the infrastructure category.

Environmental information

The final group of data relates to the biophysical environment, and are categorised hierarchically as ***Level III***. Data in this category are of a thematic nature, but are included to highlight the importance of natural resources and the environment in Africa's development. They represent naturally occurring elements and therefore constitute primary information. However, the related data sets may require some derivation or basic analysis. Specific categories such as vegetation, forest reserves, agriculture, mineral deposits, etc., are not included because these can be derived from the combination of different identified fundamental data sets.

5.4 Spatial features and attribute information

Table 8 shows key spatial features for the candidate fundamental data sets and their essential attribute information.

Table 8: Fundamental geospatial data sets, spatial features and their attributes.

Geodetic control network	Geodetic control points	Trigonometric points, coordinates	Coordinates; history of the network; network design
	Height datum	Points	Primary height values
	Geoid model	Reference ellipsoid	Name of reference ellipsoid; origin; numerical values of ellipsoid parameters
Rectified imagery	Aerial photography	Ortho photos	Date, time, scale, format, projection, level of rectification
	Satellite imagery	Orthorectified images	Date, resolution, bands, format, projection, level of rectification
Hypsography	Digital elevation model	Contour lines, bathymetry lines	Height values
	Spot heights	Spot heights	Heights of peaks
	Bathymetry	Contours, point ocean features and grid	Type, depth
Hydrography	Coastline	Coastline.	Scale, source, date
	Natural water bodies	Streams and rivers (perennial, intermittent/seasonal), canals, ponds, lakes, wetlands, wells	Unique code, name, length, surface area
Boundaries	Governmental units	International, National, provncial/ regional, district, local government, traditional authority, ward, township, tribal	Unique code, name, area
	Populated places	Capitals, urban areas, towns villages, localities, and rural settlements	Unique code, name
	Enumeration areas	EA units	Unique code
Geographic names	Place Names	Village, town, suburb, city	Unique code, name, synonyms, type, source, date
	Feature Names	River, mountain, farms, landforms, etc	Unique code, name, synonyms, type, source, date
Land management units/areas	Land Parcels/Cadastre	Land parcels	Parcel number, owner, size, date acquired
	Land Tenure		Details of ownership, vesting, and including traditional forms of land holding
	Street Address	Address point	Street number, street name, street type, postal code, place name, province
	Postal or zip code	Post code areas or zones	Unique code
	Land use planning zones	Conservation areas, heritage sites, and other restricted areas (state protected areas)	Name, area
Transportation	Roads	Main/trunk/national road, secondary road, tertiary road, minor road, street, tract	Unique code, name, surface, length, number of lanes
	Road centrelines	National, main, major, minor, trails, secondary, other	Unique code, name, surface, length
	Railways	Railway line, station	Unique code, name, type, length
	Airports and ports	Airports, airfields, landing strips, harbours	Unique code, name
Structures	Structures	Bridges, tunnels, ferries, towers, stadiums,	Unique code, name, type, latitude/longitude
Utilities and services	Power	Power stations, power lines	Unique code, name, capacity, type
	Telecommunications	Telecommunication towers, telephone network	Unique code
Natural environment	Land cover	Rangelands, forests, woodland, scrub, urban or built up areas, and wetlands	Name, surface and area
	Soils	Soil types	Name, code, area, depth, land capability, clay content, agricultural constraints, etc
	Geology	Lithogical, units/contacts	Unique Code, Name, Age, Stratigraphy
		Structure	Unique Code, Name, Type, Age
		Regional boundary Geological	Major rock formation and sequences
		Regional structure features	Type of features (e.g. faults, joint)
		Major ore deposit	Type, name, commodity

5.5 Scale and level of detail

The traditional concept of scale does not strictly apply when used in the context of digital geo-spatial data and "digital maps". Traditionally, and in present hardcopy maps, scale was, and is, chosen to allow inspection of an area of interest in a single view and on sheets of manageable size. Symbols are used on maps in order to represent phenomena and information logically and clearly, and decisions are made based on the extent of the area to be displayed on one map sheet as well as the size of the smallest feature that have to be distinguishable. Depending on the scale and purpose of the map, various classes or groups of features are represented with various degrees of detail.

In thematic maps, the principal subject is usually represented in detail by including a large range of sub-classes. However, a base map is a reduced representation of the topographic surface. As all other phenomena are shown in relation to this base, the scale of the map largely determines the amount of information that can be shown, and the amount information that can be captured and represented in a database. Specifications for geo-spatial data are therefore scale-dependent.

While real-world objects to be represented in a digital environment do not have to be scaled, mapped information is influenced by scale. Therefore, as the scale of the map changes, the map or data content also changes. In other words, the map scale determines the size of the minimum mapping area and hence the material included and/or excluded. The process of scale reduction results in generalisation, and it increases in effect progressively: the smaller the scale the greater the degree of generalisation. In the digital environment, scale becomes an issue of resolution and generalisation.

Information contained on a source map has two main components: location and meaning. Generalisation affects both. As the amount of space available for representing features on the map decreases with decreasing scale, less locational information can be given about features. Generalisation also affects the number of classes and sub-classes that can be represented. Data categorisation therefore tends to be more general in the case of large-area coverage, and starts to become more particular as the scale increases up to the local level or where the interest is more specific. In this respect the definition of a 'point' is a question of scale; a town may be represented by a point in a small-scale data set while it may be an area (represented as a polygon) at a large scale.

Data can also change their status from non-geo-spatial to geo-spatial depending on scale context. For example, malaria mortality statistics for a country might only be a single figure giving the number of deaths in the country and not be linked to a position in space. However, when seen in the context of the African continent, they can be linked to the position of countries on the continent and thus become geo-referenced. This potential of change in data type applies to a large number of statistical data.

These are important considerations in the determination of a universal set of base data for the whole of Africa. The *level of detail* of information and the corresponding application scale should therefore be indicated for data sets. The following is recommended:

Highest	Site	>1:5 000	<2.5
High	Local/municipality level	1:10 000	≤5
Medium	Sub-national/provincial level	1:50 000	≤50
Low	National level	1:250 000	≤125
General	Regional	1:1 000 000	≥1,000

Levels of detail required for a universal set of fundamental data for the whole of the continent will vary, and will range from highest to lowest. This reflects the variety of features and the range of spatial attributes that may be represented at the respective scales. For instance, road centrelines or tracts may be relevant only at large scales, and will not be shown at small scales. Similarly, the location of a well can be accurately depicted at the large scale; however, as the scale decreases wells may not be shown at all, except for very thematic purposes. The suggested scales/resolution for data from high to general level of detail is consistent with scale levels represented in the VMAP series.

5.6 Level of accuracy

When information is captured from a map the largest meaningful and acceptable scale of the information is determined by the spatial accuracy of the source (mapped) data. Using traditional hardcopy maps, positional information can be captured with an accuracy of not better than 0.1mm. This limitation defines the accuracy that data can be captured from hardcopy maps. For example, a topographic map of scale 1:50 000 cannot provide accuracies of less than 5 m. In order to increase reliability and guarantee the required accuracy of digital data it is important for the data to be captured at slightly higher accuracies.

The cut-off value of 0.1 mm in traditional maps does not exist in digital data acquisition or extraction, where zooming capabilities make it possible to measure to accuracies only limited by the hardware and software used to extract the data. This seemingly unlimited accuracy can obviously lead to incorrect assumptions regarding the accuracy of extracted digital information. It is therefore crucial that data is provided together with metadata about the accuracy of the data acquisition. This is to enable the user to judge the maximum decimal places to which the data can be considered reliable.

Table 9 gives desirable or typical accuracy estimates for the acquisition of the various data sets. Some of these are given in a 'part per million format' (ppm). These values refer to the relative accuracy between points and express the accuracy in relation to the average distance between network points. For example, if the average distance between the points of a network is 100 km, then the accuracy of distances between network points should be 0.1 millionths of 100 km, which is 0.01 m.

Table 9: Levels of accuracy

Level 0		
Geodetic control points	S & M	Zero order: 1 to 0.1 ppm; first order: 10 ppm
Height datum	S & M	0.5 to 1mm *√ K (distance between point in km)
Geodetic framework - GPS	S & M	Zero order: 5mm + 0.2ppm First order: 10mm +2ppm
Geoid model	S	0.1 m (ideal)
Level I		
Ortho-photos	M &L	Depending on scale
Ortho-images from satellite data	M	5m to 30 m
Digital elevation model		0.1 m (ideal) to 1 m
Topography	M	Depending on scale: 1m to 10 m
Natural water bodies	S,M & L	Depending on scale
Level II		
Governmental units	S &M	
Populated places	S &M	50m
Enumeration areas	L	5 -15 cm for urban areas 0.5 m to 1 m for farms
Geographical Names	n.a.	n.a.
Feature names	n.a.	n.a.
Cadastre	L	5 -15 cm for urban areas 0.5 m to 1 m for farms
Land tenure	L & M	5 -15 cm for urban areas 0.5 m to 1 m for farms
Street Address	L & M	5 -15 cm for urban areas 0.5 m to 1 m for farms
Land use planning zones	S &M	1m to 10 m
Road networks	S &M	1m to 10 m
Road centrelines	L	5 -15 cm
Railways	S &M	1m to 10 m
Power	S &M	
Telecommunication networks	S &M	1m to 10 m
Level III		
Land cover	S &M	10m to 500 m
Soil	S &M	10m to 500 m
Geology	S &M	10m to 500 m

(S = small; M = Medium; L = Large)

Accuracies in Table 9 must be understood as optimal accuracies and not as absolute criteria. In many cases these accuracies have not, or cannot, be achieved as a result of local conditions. With regard to the fundamental data sets it is recommended that levels of accuracy be consistent with requirements for mapping at the respective scales.

5.7 Metadata

The definition and adoption of fundamental data sets should promote the widespread use of geo-information, particularly with respect to data integration. However, in order to facilitate optimal use of fundamental data sets they should be widely published and understood by all users without indeterminacies or conceptual ambiguities, and there must be mechanisms in place to facilitate discovery. This, in turn, requires that there are unique definitions for all and every piece of data. Data producers and users must agree on terminology and "descriptors" of the data.

Inputs from respondents in the sub-regions suggest that the following elements are essential when capturing metadata for fundamental data sets: originator of the data set, publication date, title of the data set, format of the data set, description of the data set, purpose of the data set, date of completion, status of data set (e.g. completeness), contact details of custodian, accuracy of attributes, accuracy of spatial data, scale of maps, projection/coordinate system, datum, ellipsoid, access constraints, use constraints and distribution information and spatial boundary extent.

In addition to these, international standards for metadata should also be adhered to.

5.8 Temporal requirements

One criterion suggested by various respondents for fundamental data is that they should be of permanent value, and be persistent over time. In this respect, and by the definition in the hierarchy adopted, primary (Levels 0 and I) fundamental data sets are generally not subject to temporal variations except in terms of long-term phenomena, such as a change in the shape of a lake over time. The actual data items or parameters may be updated or changed as, for instance, in a redefinition of the geoid model or recalculation of the geoidal parameters.

Level II data sets are not primary, but depend on human decisions, and may therefore change over a relatively short time span. This also true for land cover data (Level III), but not the other candidate data in the latter group (i.e., soils and geology). Inputs from the study suggest that Level II and III data should be maintained and updated every five years.

6. Conclusions

This consolidated report describes input from the sub-regions of Africa on the issues of fundamental datasets. The results were gathered and made possible by the involvement of some of the most prominent players in the geo-spatial information arena in each sub-region. Although information was obtained from a relatively small sample of data users, these were carefully selected and adequately represent the most important and largest users of geo-spatial data in Africa. Therefore, it is strongly believed that the conclusions drawn out of this study reflect the needs and practices across Africa. Although inputs are still awaited from some of the sub-regions in Africa it is believed that they will not change significantly the results presented in this report.

This study has importantly put forward a single unambiguous definition of fundamental datasets for Africa. This was no easy task, especially in light of inconsistencies in the international literature review and the varied responses from sub-regions across Africa. The task was a necessary one to guide the process of actually determining the relevant datasets. The proposed definition reads:

Fundamental data sets are the minimum primary sets of data that cannot be derived from other data sets, and that are required to spatially represent phenomena, objects, or themes important for the realisation of economic, social, and environmental benefits consistently across Africa at the local, national, sub-regional and regional levels.

The above definition together with the concept of "candidate" data sets and their description will make the enormous task of identifying priority sets in Africa more manageable. This will facilitate the speedy capture of these datasets by appropriate mapping agencies on the continent.

In addition to the definition above, this study has identified and defined data themes that constitute the fundamental data sets. These are listed below:

1. the geodetic control network,
2. remotely sensed imagery (e.g. aerial photography and satellite imagery),
3. hypsography (e.g., contours, DEM, spot heights, etc),
4. hydrography (e.g., rivers, streams, water bodies, etc),

5. administrative boundaries (e.g., international, provincial, district, etc),
6. geographic names,
7. land management units/areas,
8. transportation,
9. utilities and services,
10. the natural environment.

These data themes are largely consistent across all the different inputs received, and with international determinations, although slightly different justifications are assigned. The practical implication is that, consistent and up to date data sets, having national coverage should be available for planning, management and decision making purposes, and that the recommended data sets are essential for all countries. Depending on the theme and scale of the data, it should be made available at the local, national, and sub-regional levels.

Importantly too is the recognition by those surveyed that fundamental datasets are not static and rigid. These data sets should and must be reviewed by the relevant role-players on a regular basis as many of them are based on priorities at that time. Priorities and development needs across Africa will and are changing and so too will fundamental datasets.

In addition to the finding above, the important and often overlooked aspect of consultation and consensus building in determining fundamental datasets is very important. Fundamental datasets at all scales are an asset to nation states, economic regions and the African continent as a whole. Once it is recognized as an asset and is seen as the bedrock from which intelligent decisions about Africa's development can be made, only then will the capture and maintenance of these data sets be regarded as a priority.

The findings of this research can be used to raise awareness of the need for the development of fundamental data sets in Africa. The recommendations outlined in this report should be used as the basis for the Mapping Africa for Africa initiative. This study is only the starting point on the long journey to have Africa comprehensively mapped.

Having now reached a common understanding and definition of what geospatial data constitute fundamental data sets, the second phase of this project will be to compile an inventory of presently available fundamental data sets in each member State. As of this writing, the Phase II survey instruments are currently in-field. The third and final phase of this project will be to develop policy, strategy, and operational plans of action to back-fill missing fundamental data sets identified from the project's second phase. Ultimately, a series of geographically exhaustive and synthetic fundamental geospatial data sets for the continent will form the keystone of African's Regional Spatial Data Infrastructure that will, in turn, complete and add considerable value to the UN Spatial Data Infrastructure, a global effort to procure, collate, vet, and disseminate geospatial data.

References

Hansen Albites, Francisco A., . INEGI, Mexico, A characterization of data in the context of SDIs, presented at GSDI7 Conference, February 2 – 6, 2004, Bangalore, India.

Gavin, E., and Gyamfi-Aidoo, J., 2001, Environmental Information Systems Development in Sub-Saharan Africa — Approaches, Lessons and Challenges, EIS-AFRICA, p. v

United Nations Development Programme and United Nations Environment Programme, (1994), International Symposium on Core Data Needs for Environmental Assessment and Sustainable Development Strategies. Volumes I and II. Bangkok, November 15-18, 1994.

Falloux, F. 1989, Land information and remote sensing for renewable resource management in Sub-Saharan Africa: A demand-driven approach, World Bank Technical Paper 108, The World Bank, Washington

World Resources Institute, 1995, Africa Data Sampler CD-ROM — A Geo-Referenced Database For All African Countries

FAO, 1998, The FAO AFRICOVER Programme, http://www.fao.org/sd/EIdirect/EIre0053.htm, accessed 3 August 2005)

FAO, Multipurpose AFRICOVER Database for Environmental Resources, http://www.africover.org/MADE.htm (accessed 3 August 2005)

KARI, 1992. Application of Geographic Information System (GIS) for efficient data storage and handling in Kenya. Proceedings of a symposium organized by Kenya Agricultural Research Institute, 5 -6 March 1992, Nairobi, Kenya.

Survey of Kenya., 2002. Proceedings of the second Kenya National Spatial Data Infrastructure Workshop. April, 2002. Nairobi, Kenya

Mavima R, Noongo EN, 2004, Spatial data infrastructures in SADC countries: Status and future directions.

Luzet, C. and Murakami, H. (eds.), Geospatial Data Development: Building data for multiple uses, "In: Nebert, Douglas D., (Ed.), 2004, Developing Spatial Data Infrastructures: The SDI Cookbook, Version 2.0"

Annex 1: Study Terms of Reference

Determination of the Fundamental Geo-spatial Datasets for Africa through User Needs Analysis

1. Introduction

1.1 The African continent is poorly mapped with little or no systematic collection and maintenance of fundamental geo-spatial datasets taking place. This is negatively impacting upon effective decision-making and development planning. Various development projects collect such datasets, but only to satisfy the minimum requirement of the project. This data collection is done in a sporadic and uncoordinated manner with no intention of maintaining such data. As a result the data becomes obsolete very soon, and also is not accessible for purposes other than for that project.

1.2 To provide the geo-spatial information required for effective and efficient decision making and development planning requires a more systematic and programmatic approach to the collection and maintenance of this information. For most development needs there are common geo-spatial information required, referred to as the fundamental geo-spatial datasets (or foundation or core datasets). These fundamental geo-spatial datasets are generally collected and maintained as part of a national mapping programme, but can also be collected and maintained at a regional level or aggregated from national to regional level. These fundamental geo-spatial datasets form part of a spatial data infrastructure (together with standards, access mechanisms and policies).

1.3 The Subcommittee on Geo-information of the Committee for Development Information (CODI-Geo) of the United Nations Economic Commission for Africa (UN-ECA) as well as other international organisations, such as the International Cartographic Association, have recognised the need to address the situation in Africa. The Mapping Africa for Africa initiative aims to address the issue of the lack of accurate, reliable and up-to-date fundamental geo-spatial datasets in Africa. As part of this initiative it is necessary to determine, from a user perspective, what makes up the fundamental geo-spatial datasets. A needs- or demand-driven approach is required to ensure the effectiveness of the collection and maintenance of these fundamental geo-spatial datasets.

1.4 Please note that socio-economic and demographic datasets, also regarded as fundamental datasets, are excluded from the scope of this contract.

2. User needs analysis

2.1 To ensure a needs/demand-driven approach to the collection and maintenance of fundamental geo-spatial datasets it is essential that a user needs analysis be performed.

2.2 The service provider must undertake such an analysis to determine :

a) What is deemed to be the fundamental geo-spatial datasets (at national and sub-regional and regional level), from the universe of geo-spatial datasets, using criteria to be agreed upon;
b) For each fundamental geo-spatial dataset, what spatial and descriptive (non-spatial) information is required to be collected and maintained, including the level of detail (spatial resolution and semantic level), accuracy and metadata.
c) Any temporal requirements to meet application needs (i.e. how up to date the dataset must be, or the time intervals between the revision of the dataset).

2.3 Information for the analysis may be surveyed from documentation analysis and postal/ telephonic questionnaires. Note that it will not be necessary to conduct contact interviews/ workshops with the users.

2.4 The service provider must include in the report the methodology and criteria used to determine the fundamental geo-spatial datasets.

2.5 The service provider must verify the findings from the user survey.

2.6 The service provider should make use of the recent study conducted by EIS-Africa / USGS on data content standards in Africa and standards published or under development by ISO/ TC211.

2.7 The service provider must indicate in their proposal the list of users that will be used in this user needs analysis. The users must be either African organisations or international organisations working in Africa for the benefit of a country or sub-region. All relevant application sectors must be covered, in particular :

- Agriculture, including food security;
- Transportation (road, rail, water and air) and communication;
- Environmental management;
- Disaster management;
- Spatial planning;
- Health;
- Safety and security;
- Water resource management and supply;
- Energy;
- Tourism;
- Housing;
- Land administration

As a minimum the following users must be included:

a) Five different national government ministries/departments (not necessarily from the same country) who are users of geo-spatial information, from each of the five sub-regions (as defined by UN-ECA) in Africa;
b) UN Environment Programme;
c) UN Development Programme;
d) UN-ECA (Division : Development Information Services)
e) World Bank.

It must be noted that the main official languages in Africa are English and French, with Portuguese in Angola and Mozambique. Documentation sent to any country must be in the official language of that country.

3. Information on datasets to be analysed

3.1 The service provider must make recommendations on :

a) What is to be regarded as a fundamental geo-spatial dataset, stating its main classification/theme, and how the dataset was agreed as a fundamental dataset;
b) For each main class/theme, to specify each component feature class (to the required level of detail) with its required spatial and descriptive attributes, metadata content and temporal requirements;
c) The number of occurrences each feature class was requested by the users; d) Definitions for each feature class, and where necessary the definition of the attributes (use data content standards).

3.2 Note : A feature-based (object-oriented) approach must be followed.

3.3 A report, in English, must be submitted, containing the above information, and the contact details of users surveyed. The report must be submitted as a MS-Word document to:

Chief Directorate : Surveys and Mapping (Attention : Mr D Clarke)
Private Bag X10
Mowbray
7705 SOUTH AFRICA
or,

e-mail : dclarke@sli.wcape.gov.za
by 20 July 2005

3.4 The report will be circulated to members of the Working Group and selected persons for comment. The tenderer will be required to address any comments and provide a revised report within three weeks of receipt of the request to make any changes.

4. Payment

4.1 All payments will be made in South African currency (ZAR) electronically to a South African registered Bank.

4.2 Progressive payments will be made based on work completed, with the final 10% of payment being made upon acceptance of the final report.

4.3 Expenses for travel and accommodation will be paid when such expenses have been incurred, based on actual expenditure incurred.

5. Tender Price

The tender price (in ZAR) must include all costs for completing the work required, with a breakdown of the main cost items. The price must include VAT. Costs of travel and accommodation must be shown separately – limits : air travel : economy class, and accommodation : 3-star accommodation.

6. Evaluation of Tender

6.1 Tenderers must include the proposed methodology to be used to carry out the work.

6.2 The provisions of the Preferential Procurement Policy Framework Act will apply. The tender will be evaluated using the 80/20 points process.

6.3 Within the allocation of the 80 points, the tender will be evaluated using the following criteria:

- 50 points : comparative tender price;
- 20 points : proposed methodology to achieve the delivered objective of the work;
- 10 points : service provider's network of contacts in the various organisations to be surveyed, which ensures that the correct persons / components provide the information for the work.

7. General

7.1 The Chief Directorate of Surveys and Mapping accepts no responsibility or liability for any loss or damage to any persons or property in the execution of the work.

7.2 The report and any supporting documentation becomes the property of the Chief Directorate of Surveys and Mapping.

7.3 The Chief Directorate of Surveys and Mapping will provide no logistical or administrative support for the execution of the work.

Annex 2: Regional Partners and Countries Allocated to them

North Africa	Centre for Environment & Development for the Arab Region and Europe - CEDARE	Algeria, Egypt, Libya, Morocco, Tunisia
Southern Africa	[SADC Regional Remote Sensing Unit]	
	Angola, Botswana, Lesotho, Madagascar, Malawi, Mozambique, Namibia, South Africa, Swaziland, Zambia, Zimbabwe	
East Africa	Regional Centre for Mapping of Resources for Development	Burundi, Eritrea, Ethiopia, Kenya, Rwanda, Somalia, Sudan, Tanzania, Uganda
Central Africa	Ministère de l'Economie Forestière, des Eaux, de la Pêche, Chargé de la Protection de la Nature	Cameroon, Central African Republic,
Democratic Republic of Congo, Equatorial Guinea, Gabon, Republic of Congo, Sao Tome & Principe		
West Africa (1)	Regional Centre for Training in Aerospace Surveys (RECTAS)	Benin, Ghana, Liberia, Nigeria, Senegal, Sierra Leone, Togo
West Africa (2)	Centre Régional AGRHYMET	Burkina Faso, Cape Verde, Chad, Côte d'Ivoire, The Gambia, Guinea Bissau, Niger, Mali, Mauritania

Annex 3: Multi-national Organisations Identified to Participate in Study

African Development Bank	No contact identified
Food and Agricultural Organisation (FAO) - GTOS Programme	Dr John Latham
Global Mapping Initiative	Mr Hiromichi Maruyama
United Nations HABITAT	Mr Eduardo Moreno
Southern African Humanitarian Information Management System (SAHIMS) Programme	Mr Georges Tadonki
Swede Survey	Mr Ake Finnstrom
United Nations Working Group on Geo Information (UNWGGI)	Mr Ergin Ataman
United Nations Development Programme (UNDP)	No contact identified
United Nations Economic Commission for Africa (UNECA)	Dr Dozie Ezigabalike
United Nations Environmental Programme (UNEP) Division of Early Warning and Assessment (DEWA), Africa Region	Mr Charles Sebukeera
World Health Organisation (WHO)	Mr Steeve Ebener
World Bank	Mr Uwe Deichmann
World Meteorological Organisation (WMO)	Dr MVK Sivakuma