USAID/Uganda: Monitoring and Evaluation Management Services Project

Report on Initial Assignment to Assess Project Status and Information Needs

USAID Contract Ref: #617-C-00-08-00012-00

Submitted by:

IBI International
2200 Clarendon Boulevard, Suite 1200
Arlington, Virginia 22201 - USA
Tel: +1-703-525-2277
www.ibi-usa.com

Prepared by:

Gwynneth M. Martin

October 2008

CONTENTS

APPENDICES

1. INTRODUCTION

1.1 Background

In May 2008, The Mitchell Group Inc. (TMG) entered into a contractual agreement with USAID/Uganda (Contract #617-C-00-08-00012-00) to provide comprehensive monitoring and evaluation support services to the Mission, serving Teams and their Implementing Partners. As described in the RFP, a key part of the Uganda Monitoring and Evaluation Management Services (UMEMS) activity is to provide USAID/Uganda and its Implementing Partners (IP) with a web-based reporting system that interfaces successfully with the Geographic Information System (GIS) recently developed for the Mission.

Geographic information systems provide capabilities for management of spatial data, in this case data relating to geographic locations on the earth's surface, enabling the production of maps to present information in a manner which will assist with interpretation and communication. This will allow examination of Mission program data with respect to its spatial distribution, proximity and association with various geographical features, and provide the basis for analysis of the performance data in a spatial context, hence the ability to compare districts, counties, etc.

The spatial component of the planned UMEMS system development can be broken down into the following elements:

1. The existing GIS – hardware and software, etc - comprising a spatial data management system
2. The GIS spatial database (including map layers such as roads, rivers, topography, political boundaries, infrastructure, population, etc.)
3. Protocols for the acquisition of point-location data (to provide the spatial dimension as required)
4. A web-based data collection and management system (input of performance data by USAID/Uganda and IPs, and the storage and management of the data)
5. Conversion and linkage of performance data with the appropriate geographic layer
6. Analysis of the performance data in a spatial context.

1.2 Approach

A review of elements 1 and 2 above is a first essential step to determine exactly what is currently in place and its operational status, and to evaluate its appropriateness for the needs of the Performance Reporting system being planned. For elements 3, 4 and 5, detailed user requirements for the management (acquisition, storage, retrieval and display) of the spatial data and its relationships to non-spatial performance data need to be established as the basis for system design and implementation. Step 6, the development of appropriate analysis capability and applications, will evolve as the basic operational system is developed and implemented.

A two-week consulting assignment was organized to begin addressing the spatial component of the UMEMS Performance Reporting System. The terms of reference are as follows:

- *Review the Mission's GIS database with a view to assessing its suitability for storing and manipulating the data envisaged by the Mission;*
- *Interact with the staff of the Mission's Program & Policy Development Office and the Mission Teams to determine their expectations of the GIS system to be developed;*
- *Discuss with the TMG Database Consultant the needed integration between the proposed system and the web-based databases into which Mission partners will, on a quarterly basis, enter their quantitative performance data pertaining to performance indicators;*
- *Prepare a detailed work plan and timeline for the way forward.*

The assignment was planned to coincide for the first week with an assignment for the TMG Database Consultant, Mr Niyi Fajemidupe, to define the requirements for the web-based quarterly reporting system. This interaction will facilitate planning the needed integration of the spatial elements with the indicator-related information and will begin the process of ensuring compatible linkages between the spatial and non-spatial data with consistent standards for quality assurance.

The overall approach involved meeting with staff of the Mission to discuss the existing GIS system and to establish their needs and expectations of the planned Performance Reporting System. The early meetings, involving the TMG Database Consultant, addressed the system as a whole. Subsequent meetings were organized with the Mission, with Ugandan Ministries, and with Makerere University to follow-up and determine detail of requirements for the management of spatial data.

2. GEOGRAPHIC INFORMATION SYSTEMS

2.1 Introduction

The functionality of a GIS parallels that of a DBMS, i.e. data input, storage, retrieval, reporting and so on, but with the additional facility that a GIS manages **spatial** data. These data are about objects or phenomena that are associated with a position, commonly a location on the surface of the Earth, and are often best visualized by producing map output. Thus GIS are designed to handle coordinate data defining positions of points, lines and areas, to be able to encapsulate spatial relationships in the data structure, and to produce graphic (map) outputs. A GIS also manages, and integrates in the system, what is termed "attribute data". These are non-spatial items that are associated with the spatial feature. For example:

- a well is a spatial feature (a point) whose position is given by a pair of co-ordinates (latitude and longitude); it may have attributes of "owner", "depth", "water quality", etc
- a road is a line feature (a series of pairs of coordinates) which may have an attribute of "class", "surface material", etc
- a district is a closed area (polygon) which may have attributes "name", "population", and other socio-economic data.

There are multiple definitions and interpretations of what is meant by the term "geographic information system". It is variously described as software, as a database, as some combination of hardware, software, database and procedures. A fully operational system will involve hardware, software, a database, and procedures, as well as **people** who are knowledgeable in the technology, have an understanding of the data, and the abilities to carry out the procedures. Section 2.2 below describes the status of the GIS in the Mission under those 5 headings.

As shown in Figure 1, a GIS typically has two major software components: a Relational Database Management System (RDBMS) that holds and supports the processing of the attribute data, and a Spatial Data Engine – a specialized software package that deals with the spatial linkages and processing. These may be fully integrated, but often are linked through standard protocols.

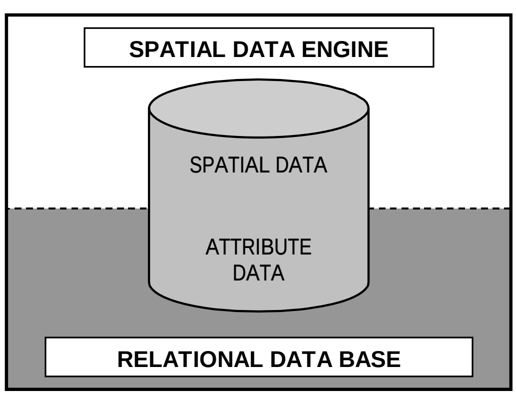

Figure 1: Geographic Information System

The implementation of an operational GIS is a significant undertaking. Often the first (and simplest) steps are the acquisition of hardware and software with or without consideration of the particular functionality and capacity required. A common approach is to utilize a generic or general purpose GIS that can be customized to the application requirements. The design of the database and building of procedures, both automated and manual, can only be done effectively through analysis of those requirements – who are the users, what are the outputs required, where does the data come from and how, definition of data items (type, units, form, validation rules, etc), frequency of various operations and so on. Effectively everything about what the system will do needs to be specified – in detail – and properties of data items and their relationships must be carefully defined. This is similar to the development of a regular database application, with the additional need of documenting all the relationships between the spatial elements.

2.2 USAID/Uganda GIS Status

2.2.1 Hardware

The Mission operates an internal network supporting approximately 60 people, with 3 IT staff managing and maintaining the network. They do not anticipate any problem in providing processing capacity and storage space, although no estimates for volumes have been given.

Currently the only hard copy output that can be produced is from A4 printers. A larger format is commonly required for map output, and color printing is highly desirable. An A3-sized printer for map output would likely be sufficient and the occasional need for larger formats could be accommodated through commercial suppliers.

2.2.2 Software

There are two pieces of extant relevant software, each of which has been referred to as the USAID/Uganda GIS system. These are:

ArcGIS

- a commercial GIS package from Environmental Systems Research Institute, ESRI – a US company and the predominant GIS vendor internationally
- commonly used, "standard"
- software licenses in place for 3 modules
- not yet used by anyone in Mission.

ESRI markets a suite of software modules. The Mission's licenses provide functionality for desktop mapping and basic spatial analysis. This includes managing layers of spatial data, composing, displaying and printing maps, linking to and integrating tables of attribute data, and analyzing and visualizing the resulting spatial relationships. The map composition functionality is very extensive and allows for a broad range of options for coloring and symbolization, annotation, north arrows, scale bars, etc. Users can zoom and pan across maps and interrogate the data by point-and-click.

The capabilities for spatial analysis with linked attribute tables is also provided –

allowing for such procedures as colouring of area according to summary attributes or combinations (averages, min/max, ratios, etc) and for symbolization through pie-charts, bar-graphs, as well as adjacency and proximity analysis.

The software has been successfully used in a wide range of applications in many different countries. The broad range of functionality however, implies a level of complexity and hence in many cases technical experts will develop user-friendly procedures to enable "non-experts" to use the system for commonly required functions.

USAID GIS DB v1.0

- application package developed under contract by Makerere University Geo-Graphical Services – personnel from Makerere University Institute for Environment and Natural Resources, and Elite Technologies (IT solutions company)
- package intended to "assist the USAID/Uganda Mission in better planning, monitoring and presenting its program"
- developed using GIS Open Source software (for mapping), MS SQL Server (for database management), Visual Basic (for processing logic) and Visual Studio (for user interfaces)
- not yet used by anyone in Mission
- includes some base layers of spatial data e.g. administrative boundaries, rivers, roads.

The contract was awarded in April 2007 and the Statement of Work outlined the usual steps in system implementation i.e. establishing detailed specifications, system design and development, installation and testing, and training. However, when the completed system was demonstrated, additional capabilities were requested – capabilities which had not been in the original scope of work – and these necessitated re-design and further development. This work is now complete.

Unfortunately there was a problem with access to the software in the Mission, so it could not be used hands-on. However, on October 7th the developer made a presentation to the COP, the two visiting consultants and Mission staff, demonstrating the system by walking through the various menu-driven functions. The system includes a "help" facility (based on Windows Help) and that was used as a basis for summarizing what the system was designed to do (see Appendix I).

At a later point in this assignment, the developer provided technical documentation by way of an Enhanced Entity Relationship Diagram and database schema. These were useful to confirm and provide detail on the Mission requirements in some areas.

2.2.3 Database

There is no existing spatial database. However there are several spatial datasets – in fact, potentially duplicate copies of the same dataset. (The distinction drawn between a data**set** and a data**base** is that the latter is intended to be available to multiple users and will have a well-defined structure with relationships specified, and explicit consistent standards and specifications for data items.)

2.2.4 Procedures

No procedures have been developed.

2.2.5 People

A GIS Specialist post has been created and staffed. The incumbent, Fortunate Muyambi, took up his duties October 14th i.e. mid-way in the period of this consulting assignment. His background is in natural resource conservation and he gained his initial expertise in GIS in that context. Over the past five years, he has held positions involving the application of GIS in Uganda Wildlife Authority, the US-based Ugandan Mountain Gorilla Conservation Fund, the Ministry of Water and Environment and, most recently, the Electrical Distribution Company. He is experienced in using ESRI's ArcGIS software and has attended their User Conferences – in fact, won an award for a paper entitled "GIS Applications in Utility Companies", delivered at the 2008 Conference.

On Friday October 17th, at a meeting in the Mission, he reported that he had installed the ESRI software successfully and was collecting base layers for review and validation.

2.3 Conclusions

There is not an operational GIS in the Mission presently.

The post of GIS Specialist has just been filled. This is one essential first step and an essential component of a system. His workplan and associated milestones need to be coordinated with the overall workplan for development of the Performance Monitoring System (see Section 6).

With respect to the core "GIS" elements of the two software products listed above, ArcGIS is a generic, commercial off-the-shelf package while the Makerere system uses open-source (free) software. Both have the some of the same base functionality but the former is potentially more powerful and robust. There is also considerable support provided by the vendor of ArcGIS and through numerous user groups; the level of support for an open-source product is unpredictable. Similarly since use of ESRI products is widespread, there is considerable consulting expertise available if needed; this is less certain for any open-source package. Since the Mission has already acquired licenses, the cost is not an issue. The Mission appears inclined towards adopting ArcGIS and the 2006 Needs Assessment report[1] recommends use of ESRI products with the implication that they are a *de facto* standard for use in the Agency. It is positive as well that Mr Muyambi has used ArcGIS in his previous posts and is conversant with the product so there is no learning curve.

Considering these points, ArcGIS is the preferred choice for the implementation of the Performance Reporting System.

An added factor is that the Makerere system provides more than GIS elements; it includes a relational database developed to manage indicator and other related data i.e. duplicating some of the data and functions that the indicator-related part of the

[1] The existence of this preliminary GIS Needs Assessment was made known late in the course of this assignment. The general recommendations made at that time (regarding base layers, metadata, standards and so on) are in line with those made in this report.

UMEMS system development is intended to address. The scope of the latter is broader and the system is intended to be web-based.

However, the work done in the custom development is not wasted. It has raised awareness (and expectations!), and examination of the system and associated documentation is useful in giving a view of some user requirements and the data to be acquired and managed. In particular, the database designer (Emmanuel Mnzava of Elite Technologies) was very helpful in clarifying some points. Should the requirement for local system development resources arise, Elite and MUIENR would be a good source since the staff already have some knowledge of the Mission requirements and understand the need to use both RDBMS and GIS.

3. CONCEPTUAL OVERVIEW OF THE SYSTEM

In documents and discussions to date, there are references to a Quarterly Reporting System (QRS) and to a Geographic Information System (GIS). The QRS was envisaged to handle primarily indicator-related data and initial discussions in the Mission indicate that it will be required to produce not just quarterly reports but also other types of outputs. It will use a standard relational database management system (RDBMS) software package. There is also a requirement to manage non-indicator related data in the RDBMS – “other performance data” - that is not essential for quarterly reporting but that has a spatial element and is required for GIS processing.

Essentially the end-product should be a unified system as shown in Figure 2 and referred to as the “Performance Reporting System”. It will give facilities to:

- acquire performance-related data from Mission staff and Implementing Partners (IPs)
- manage the data to ensure consistency, quality and availability
- produce reports, both standard and ad-hoc
- visualize the data with map outputs.

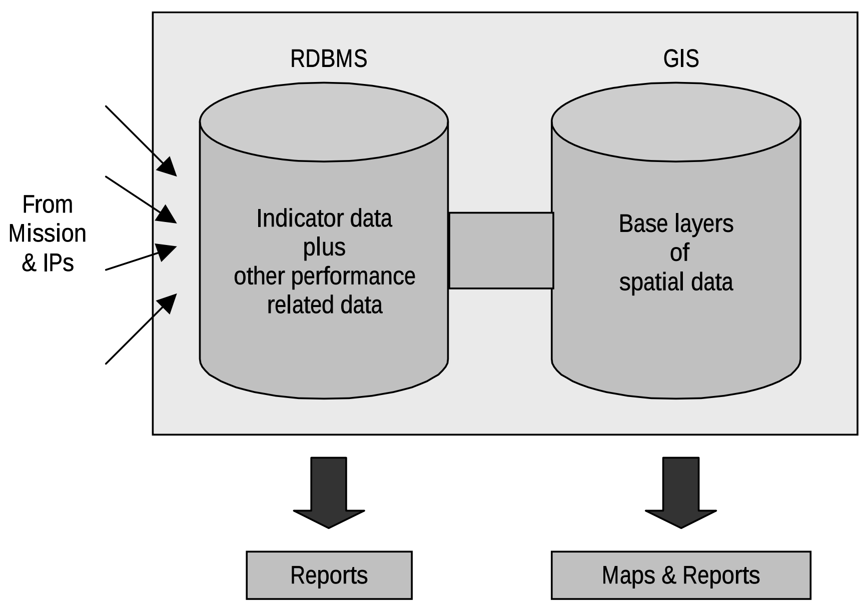

Figure 2: Performance Reporting System

As noted in the previous section, the term GIS usually refers to software with functionality for both spatial and non-spatial data management and the latter is often in the form of an RDBMS. Thus in this case, the GIS (rhs of figure) does have capability to handle attribute data, but the bulk of the non-spatial data is handled in the RDBMS (lhs of figure). The integration of the two is discussed in Section 5.

A first step in system design is to understand and document user needs. This involves defining the data required, identifying the sources of those data, input and validation requirements, the types of output required, how the data will be maintained, and so on. The requirements and specifications for the management of indicator-related data are being addressed by the consultant Database Specialist. This report is concerned with the requirements related to the spatial data – both the base layers and the other performance data. However there are several points where the two sets of requirements must come together. For instance:

- All performance data, indicator-related and other, is managed in the RDBMS and so must be included in the database design. This includes table definitions, input forms, naming and coding standards, validation rules, update mechanisms, etc.
- The indicator-related data are not spatial *per se* but they are anticipated to be "by district" in many cases (as well as potentially by county and sub-county) and hence mapped outputs will be essential. The administrative area names, codes and so on must be standardized and made consistent so that there is an unambiguous linkage of the indicator related items (attribute) to the spatial base layer. This is a very essential point – all databases that will be linked must use the exact same spelling for administrative area names or a consistent set of codes or abbreviations, and these must be made available to and used by all providers of data.
- A similar concern applies to any qualitative performance data – that uses grade assessments (good-better-best, increasing-static-decreasing, etc.). These must be standardized in application and coding.

4. SPATIAL INFORMATION REQUIREMENTS

4.1 Description of Spatial Data

As shown in Figure 2 above, the spatial data in the Performance Reporting system consists of two parts - base layers of spatial data managed within the GIS, and other performance related data managed within the RDBMS. These two are "spatial" in the sense that they include coordinate values i.e. latitude and longitude pairs, defining the position of features on the earth's surface. Also in the RDBMS are indicator-related data that are not spatial *per se* (there are no coordinate values with these data) but they do relate to administrative areas and need to be linked to the spatial definition of those areas for mapping. This linkage is essential to produce the required map outputs (see Section 4.5).

4.2 Base Data Layers

In a GIS, data is organized into layers, each layer representing a set of like features. The features may be points, lines or polygons defined by coordinates and attributes as described in section 2.1. The data layers are used singly and in combination to analyze relationships and answer spatial questions – what districts have shown the most/least progress; how is that related to investment, how is it related to socio-economic factors; which activities have been most successful in which geographic areas, and so on.

A first priority for the Mission is to assemble a base of good quality layers of spatial data to be available for use in producing maps. Many of the layers needed are available in digital form from responsible Ugandan Ministries - in fact, some have already been acquired by Mission staff. By the end of this assignment in Uganda, Mr Muyambi had in fact begun the assembly of base layers.

The primary requirement for the performance management system is to have base layers of Uganda's administrative areas since much of the data planned to be held in the RDBMS is related to districts, counties, sub-counties, etc and could be presented or "visualized" in map form with the appropriate base layer data. Other layers of potential use are lakes, rivers, roads, and other infrastructure – these would be useful to include on maps to increase user familiarity and orientation.

The following describes these specific layers with notes on any issues and questions.

Administrative Areas

- available from Uganda Bureau of Statistics (UBOS)
- organized hierarchically i.e. region-district-county-subcounty-parish-village
- main requirement is for district-county-subcounty levels;
    - there are references to 4 regions (North, South, East, West) but it is unclear how well-defined the boundaries are; the Mission also uses a number of regions for specific roll-ups
    - parish and village are below the level commonly used with indicator data

- each administrative area has a unique code[2] and full name; details of these need to be specified for incorporation and use in the RDBMS; these must be consistent and are essential for linking the indicator data with the base spatial layer
- tables of the hierarchical structure, i.e. which counties are in which district, etc., are needed for use in the RDBMS e.g. to build drop-down lists in data-entry screens
- UBOS provides attribute data (demographic items) with the administrative layers; definitions of these are needed to avoid misuse and misinterpretation and to ascertain how they might effectively be used with indicator data. [There is great potential to apply the demographic and other socio-economic data in conjunction with the monitoring indicators to analyze progress and assess relativity of conditions. These would be the more "complex" types of outputs referred to in Section 4.5 below.]
- districts have been redefined in the past (there were approximately 40 in the mid-1990s and are currently 80); this is done by re-allocating existing parishes to new districts; should this happen in the future, the nature and timing of required updates needs to be addressed. [Although the practice has very recently been forbidden, it is unclear whether this edict has been made law. Consequently the capacity to update the administrative boundaries and their spatial relationships may be required.]

Hydrology

- available from the Surveys and Mapping Department but unclear how current the data are; possibly originated from SPOT imagery in the mid-1990's with inconsistent updating
- basic features required are lakes and rivers to be included on map output products as added context i.e. only needed for visualization
- definitions of the attributes of rivers are needed e.g. to use in symbolization - thickness of line depicting river could be varied depending upon the value of an attribute giving the classification of the river.

Roads

- available from the Forestry Authority and understood to be derived from up-to-date imagery and ground-truthing; unclear whether there is on-going maintenance
- as rivers above, needed for visualization, and (possibly) for planning future programs
- definitions of attributes of roads are needed (particularly those related to the relative carrying capacity, as indicated by road surface and width, or lanes).

These base layers are among those which the mission has begun assembling. The intent is to have the layers organized in a unified fashion, with full documentation of content,

[2] In relation to codes for district and county, there is reference to "two UBOS systems". This is to be clarified.

currency etc. It is anticipated that the questions and issues raised above will be resolved to a large extent in this process. Ongoing contact with Mission staff is planned in order to track changes and improvements as this work progresses (see Section 6).

4.3 Performance related data (not indicator-related)

4.3.1 Anticipated requirements

Initial discussions with Mission staff, and review of both the MUIENR system and the 2006 Needs Assessment indicate that these data primarily relate to a point on the earth's surface e.g. the location of a school, or health facility, or borehole, and so on. Obviously the kind of point it is depends upon the nature of the project being undertaken and reported on. All points will need to be defined by a coordinate pair (latitude and longitude) but the other (attribute) data associated with the point will vary depending upon the nature of what is at the location. For example, for a school an attribute may be "type" i.e. whether it is primary, secondary and so on, and for a borehole, an attribute might be depth.

With respect to determining what types of points should be included and defining their attributes, there is a trade-off between attempting "open-endedness" and efficient delivery of functionality. To allow for the former, providing in an operational system the ability for a user to add other types of points, would involve "standardizing" the attributes of the points in some way. It would be more efficient to invest in determining up-front the types of points to be included. This should be done from the perspective of selecting those data that are clearly of use in contributing to the goal of monitoring project performance.

Requirements for three types of point were identified during this assignment, namely schools, health facilities and boreholes. Descriptions of these and their associated attributes are given in Appendix II. There are also descriptions of attribute requirements in connection with performance data related to malaria, some of which relate to health facilities i.e. to points, but some of which relate to a sub-county. These latter, as with district indicator-related data, need to be used with coordinate data defining the administrative boundaries in the base layer (see Section 4.5).

It must be stressed that these descriptions are only a preliminary starting point. They should be reviewed and refined to determine:

- whether there are more "types" of points to be added
- that the attributes are described correctly
- if more attributes need to be included
- whether there are more complex relationships between attributes i.e. many-to-many relationships.

All these details will form the basis for specifying exactly which data items will be held in the RDBMS, their relationships and how they will be managed.

Another question that has arisen in discussion of requirements was the timing of reporting this performance related data. The indicator-related data is collected quarterly and it is anticipated that input to the RDBMS will follow that path. It is desirable that non-indicator related data would be collected on the same basis i.e.

values provided for each quarter. This would allow a unified picture to be obtained and these could be rolled up to semi-annual and annual totals, used for quarter-by-quarter comparisons, etc. However, in most cases there is currently no commitment from IPs to deliver non-indicator data on a quarterly schedule. Without that, logical integration of the data to produce complete and meaningful outputs will be very complex.

4.3.2 Related Sources of Data

As detailed in the previous Section, the requirement to manage data on schools and health facilities was identified as a priority, to enable production of maps showing the location of these with some display of associated attribute values (see Section 4.5). Existing data may be relevant for baseline purposes i.e. monitoring change, and some possible sources include the following:

Ministry of Education - school data

- Africon's project undertaken in 2006 to assemble a digital database of all school locations in Uganda
- at completion of project, insufficient available resources to sustain the work i.e. no updating of information has been done
- unclear whether data is now available
- the geographic location of each school was recorded with attribute information such as category of school, pupil/teacher ratio, gross enrolment ratio, etc
- an MOU for the development of an Education Management Information System has recently been signed; there is apparently significant funding from Microsoft Corporation in connection with this development.

Ministry of Health - health facility data

- there is an inventory (hardcopy) of health facilities, dated 2006
- not determined whether there are any procedures or plans to keep this current
- the inventory has no coordinate data; health facilities are listed by parish (and therefore by district and subcounty)
- attributes recorded are the name of the facility, its owner, level and status
- although there are seven levels of facility (I to VII), the inventory:
    - does not include level I (the "village" level)
    - groups levels V, VI and VII under the single description "hospital".

Although it is not clear whether the data may be readily available from either of the Ministries, or how complete the datasets may be, it would be extremely valuable to determine any standards used for naming/coding the schools or health facilities. Further, if standards are evolving, then these should be taken into account for potential use in the Performance Reporting System and the project could provide useful benefits to the Ministries by updating and improving the datasets.

4.4 Data Input and Validation

Base Data Layers

These are available and in common use by organizations and on projects across Uganda although the currency of some of the data may be in question. As has been noted in Section 4.2, Mr Muyambi has embarked upon the task of assembling base layers in the Mission GIS, identifying deficiencies and addressing questions of data definition, coding used, currency, etc

Indicator-related data

The input of indicator-related data into the RDBMS and validation of those items is under the purview of the Database Consultant. However, as noted in Section 4.2 there is a requirement for the names and coding conventions established in the base layer of administrative boundaries to be incorporated into the RDBMS. These would be used in validation e.g. to populate drop-down lists of district names, to ensure correct link of county to the district in which it lies, etc. It is also essential for mapping that the administrative area names be totally consistent with those used in the base layer.

Performance related data (not indicator-related)

It is anticipated that these data would be input to the RDBMS by Implementing Partners through a user-friendly interface, much as the indicator-related data. Since the descriptions of these data are only preliminary at this time (see Appendix II), detailed specifications are not possible. However, in general, entry of data relating to points will involve:

- specifying the IP name/identifier (using drop-down list)
- specifying the district (using drop-down list)
- specifying the fiscal year/quarter being reported[3]
- entering the geographic coordinates of the point
- entering other parameters dependent upon the type of point (using drop-down lists of allowable values where possible).

The use of drop-down lists validates data entry to a large extent. Validation of the geographic coordinates is more problematic. Ranges of allowable values can be set but validation is best done by visual inspection of a map with district outlines and the points posted i.e. using the GIS. This would verify that the points did in fact lie within the specified district.

4.5 Outputs Required

This report addresses requirements relating to spatial data in the Performance Management system and the type of outputs discussed in this section are cartographic visualizations i.e. maps. These will allow for examination of Mission performance data with the perspective of spatial distribution.

[3] As noted in 4.3.1, the issue of quarterly reporting remains to be addressed.

There are two general types of map output required:

- those concerned with indicator values which have been entered at the district level
- those concerned with the non-indicator related performance data.

Examples of these two types are described below.

Indicator values by district

- the districts (polygons) are defined in the base layer of administrative areas
- indicator values in the RDBMS are organized by district i.e. effectively a tabulation of districts with the corresponding indicator value(s)
- a variety of types of maps can be constructed using the GIS by linking the indicator values to the appropriate district polygon, for example:
    - simply posting the values of the indicator in the district polygon
    - grouping the value of the indicator into classes and colouring the polygon according to its class
    - displaying pie-charts, histograms, etc at each district
    - using information from the base layer with the indicator value to compute other values to be mapped e.g. computing values per capita, per hectare, etc.

As has been said in other parts of this report, the production of these is only possible if the indicator values can be linked to the spatial definition by using identical naming/coding systems for the districts.

Other performance data - point types

- the coordinates of the points are defined in the RDBMS with associated attribute values
- again, a variety of maps can be produced using the GIS:
    - simply posting the location of the point (school or health facility etc)
    - using different symbols, colours etc to show attribute values
    - highlighting best performing facilities and/or areas of concern.

Mapping of non-point data, for instance relating to a sub-county, would be similar to that described for the district indicator values above.

Maps may be generated for the whole country or for a specific district. Further it is possible in the GIS to define any arbitrary area and select only that area to be shown on the map. In addition, any data from base layers existing in the Mission GIS may be selected for inclusion on the map e.g. displaying rivers or major roads.

5. INTEGRATION

As shown in Figure 2 (page 8), the two components of the Performance Reporting System – the RDBMS and the GIS – are to be linked. Current plans are that they will be physically remote from each other for at least the next two years. The RDBMS will be implemented and operate from a server in the UMEMS office and will be web-accessible. The ESRI software modules constituting the core of the GIS, are installed on the Mission's internal computer network i.e. no remote access is planned.

The indicator data and other performance related data in the RDBMS need to be accessible to the GIS to produce the required outputs. Ideally this could be done by making the link between the two "transparent" – the RDBMS data can be "seen" and used by the GIS. In theory, ArcGIS has the capability to "connect with remote databases" (quote from the official ESRI website) but this will likely require the acquisition of software modules to augment those currently in place. Exactly what would be needed is to be determined along with the associated costs. The additional level of complexity, and the implications on resources and potential required capacity building also need to be considered.

An alternative, and pragmatic approach, is to set up mechanisms to download the data from the RDBMS as required. This is much simpler technically but has the disadvantage that the data has now been copied, and so the GIS would be using a snapshot of what was current at that point in time. There is always the danger that the "copy" is used subsequently when the authoritative database has been updated. There would need to be procedures established to prevent this. However, it may be advisable to adopt downloading of data as an initial approach while testing the operational functionality of the RDBMS, ensuring the spatial connections are made, and verifying the form of the desired outputs.

Also, in the longer term, the RDBMS will be operated from within the Mission, and connectivity between the two components may then be simpler.

6. PROPOSED WORKPLAN

This first assignment was very much by way of an assessment of status and needs. Two factors are now apparent – first, the establishment of a GIS in the Mission is at its inception stage and second, many details regarding the requirements for management of what is referred to as "other performance data" remain to be determined. Consequently the timeframe in the notional workplan that had been drawn up previously may need to be extended. For instance, a significant amount of GIS related training was anticipated in the first year and the timing may be more appropriate early in year 2, after the data and requirements are better defined and further development completed.

The proposed workplan needs to be coordinated with:

- the activities of the Mission GIS Specialist, Mr Muyambi
- the workplan currently being established by the Database Consultant, Mr Fajemidupe.

At this point, these have both been discussed but remain to be verified.

Regarding the Mission GIS, Mr Muyambi intends to establish base layers of data in the system, including obtaining and reviewing both metadata and associated attribute values. He expects to make substantial progress by the end of November. The administrative areas layer, which is key to the Performance Management system, is recognized as a priority for the Mission as a whole so will be included in the first tranch of layers to be established. In addition, there was discussion of Mr Muyambi undertaking some type of demonstration project by assembling a limited set of existing non-indicator performance related data and producing some examples of potentially useful map outputs. This would assist in confirming with users the types of output required and aid in defining the performance data items in more detail, as well as helping to estimate the effort required to put in place an operational system. Details of exactly what would be included in such a demonstration and the anticipated timeframe remain to be confirmed. Hopefully some results might be forthcoming by the end of December, but that depends upon the ease with which relevant test data may be assembled, and on the possibility of other demands on Mr Muyambi's time.

Regarding the work of Mr Fajemidupe, a requirements statement covering the management of the indicator-related data is to be delivered this week (i.e. week ending 24th October). As discussed during this assignment, the proposed timeframe for development is to establish an operational "Version 1" of the RDBMS by January 1st 2009. This would enable IPs to use the system to report the results of the quarter Oct 1-Dec 31, entering their data over the period Jan 1-31. None of the requirements related to the other performance data would be addressed in this first system rollout.

Given the above, contact is to be maintained with both Mr Muyambi and Mr Fajemidupe and the following tasks undertaken:

- review metadata and attribute values for the administrative areas base layer in the mission GIS to assess potential for use
- obtain the specification for any relevant demonstration project and review results and user feedback to determine implications on requirements for the Performance Reporting System

- review the requirements statement for indicator-related data from the viewpoint of spatial requirements
- ensure that the rollout includes names and codes for administrative areas that correspond to those established in the base layers of the Mission GIS.

Following up questions raised during this assignment, another task to be undertaken is to determine the extent of the functionality of the ESRI products currently in place in the Mission regarding connection to other databases.

The anticipated timeframe for all of the above is Nov 1-Jan 31 and all work would be with oversight of the COP. The estimated level of effort is not more than 15 days.

Subsequently a second assignment to Kampala is required to undertake the following:

1. Confirm the progress made in establishing the Mission GIS.
2. Address the detailed requirements for other performance management data:
   - confirm the types of data to be included in the system
   - develop specifications for content and structure of the data, for input and validation procedures, and for retrieval requirements
   - establish with the Database Consultant a development plan to include these requirements in the RDBMS.
3. Review the status of the RDBMS:
   - with a view to expanding the scope to include the types of data defined above
   - to examine the indicator-related data input by IPs
   - to work with the GIS Specialist to produce sample maps of indicator-related data from the RDBMS.
4. Assess the need for workshops and training.

The estimated level of effort is 20 days.

APPENDIX I

USAID GIS DB v1.0 – Summary of Functionality

As noted in the body of this report, the following summary of what the MUIENR system does is based on the content of the "help" facility (file provided by the consultant) because there was a problem in the Mission accessing the system in the Mission to use it hands-on.

Since the system was developed in consultation with staff of USAID/Uganda its functionality should reflect at least part of the requirements for handling spatial data.

- System administration allows different levels of "user"
- "Projects" are created, may be saved, re-used; appear to be the basis for producing maps; not totally clear what constitutes a project
- Management functions are provided for the following data entities:
  - framework: allows for objective, program area, element and sub-element with indicators at any level
  - projects
  - points: with possible attributes of type, name, x and y coordinates, up to five attributes ("classification values")
  - location (appears to refer to administrative areas).
- A query tool is provided; for point data, a user interface can be used to build an SQL-type expression; import and export to and from Excel is possible; also backup and restore of individual tables.
- Data presentation allows the production of "success stories" and a number of reports (format apparently pre-defined).
- GIS features allow for selection of data, zoom, pan, etc, the building of legends and so on.

APPENDIX II

Preliminary Data Descriptions

The performance data values reflect results achieved over a period of time and so each of the following types of data must include an attribute "time period covered". There are issues relating to this that remain to be resolved.

School point data

- there must be a unique identifier for each school with the coordinates of that point
- district and school name would be recorded and is possibly a unique combination
- attributes also include:
  - responsible IP
  - type of school (primary, secondary, …)
  - type of USAID-funded work undertaken (constructed or rehabilitated)
  - # students (disaggregated by sex)

Health facility point data

- there must be a unique identifier for each health facility with the coordinates of that point
- current usage in the Ministry of Health indicates that a health facility has a unique name within a district i.e. a combination of district + name is unique
- attributes also include:
  - responsible IP
  - parish
  - class of facility (Health Centre Levels I to VII)
  - owner (government, NGO or Private)
  - status (functional, under construction, for upgrading, complete)
- there is a non-digital inventory of health facilities

Malaria point data

- these data relate to health facilities as described above i.e. use a unique identifier for the facility and the coordinates of that location
- the attributes depend on intervention type - IPT or SET:

  for IPT, attributes also include:
  - responsible IP
  - # pregnant women at clinic
  - # pregnant women received treatment

  for SET, attributes also include:
  - responsible IP
  - # treatments prescribed

Borehole point data

- there must be a unique identifier for each borehole with the coordinates of that location
- attributes are to be determined

Malaria subcounty-level data

- the attributes depend upon intervention type - ITN, IRS or HBMF:

  for ITN, attributes include:
  - responsible IP
  - subcounty
  - # nets distributed to pregnant women
  - # nets distributed to children under 5

  For IRS, attributes include:
  - responsible IP
  - subcounty
  - # houses sprayed
  - # pregnant women protected
  - # children under 5 protected
  - total # people protected i.e. total in households

  for HBMF, attributes include:
  - responsible IP
  - subcounty
  - # children received treatment