

# **Benefits of a Relational Database Management System (RDBMS) for the Geothermal Resources Department (GRD) of Uganda**

**Peter Mawejje<sup>1</sup>, Jacinta Achieng<sup>1</sup>, Jacqueline Nnakirijja<sup>1</sup>, Jeffrey Benegar<sup>2</sup> and Robert Kennedy<sup>2</sup>**

<sup>1</sup>**Geothermal Resources Department, Directorate of Geological Survey and Mines, Ministry of Energy and Mineral Development, Entebbe, Uganda**

<sup>2</sup>**Tetra Tech, Inc., United States**

## **Keywords**

*Decision support, geothermal data management, relational database, Uganda*

## **ABSTRACT**

A great deal of data has been collected from the different geothermal prospects of Uganda. Much of it is disassociated and isolated (“siloeed”) among individuals and departments. The exploration and development of geothermal resources not only generates data, but successful completion of these activities depends on both validated data, as well as integrated and secure access to it. These data span a broad spectrum of disciplines, e.g., geology, geochemistry, geophysics and reservoir engineering, GIS and remote sensing. As Uganda refines her conceptual resources models, the different technical personnel must be “all-rounders” able to understand the language of the multiple disciplines and data types involved and correctly interpret the diverse data. A modern RDBMS is the best way of integrating this vast amount of disparate information in order to correctly make important decisions, such as where to commence exploratory drilling.

In order that Uganda fully develop her indigenous geothermal resources, GRD, with assistance from the “EAGER” program of the UK’s Department for International Development, is constructing just such a geothermal RDBMS to enable users with different devices to access information from a variety of sources, e.g., field work, geospatial information, and legacy reports. The system will physically reside in a secure, central location, its stored data subject to rigorous quality assurance screening, yet providing true simultaneous access. It will facilitate disseminating technical information to stakeholders on demand, as well as its evaluation by developers to apply for multilateral financing.

## 1. Introduction

Uganda has an unsustainable energy mix that greatly hinders her economic development. The desire is to transform from a largely rural, subsistence-agriculture-based economy to an urban industrial economy. Currently, 84% of the population relies on biomass energy (either fuelwood harvested directly from the forest, or charcoal inefficiently made from that) because they do not have access to electricity. Development of Uganda's geothermal resources is one of the key proposed strategies to diversify Uganda's energy mix and facilitate Uganda's economic transition sustainably, through increased electrification of the country (National Planning Authority, 2007). Developing geothermal energy will greatly boost several already existing industries such as horticulture, tourism and fish farming. The ripple effect from these activities means increased household incomes as several studies show a direct relationship between the availability of electricity and quality of life as well as social stability.

Currently, Uganda has a Geothermal Resources Department (GRD) under the Ministry of Energy and Mineral Development (MEMD), which is mandated to exploit, develop, regulate and execute complete supervision and control over the country's geothermal resources. Despite ongoing efforts, geothermal exploration in Uganda has not succeeded yet in deploying utilizable resources. This can be attributed to several obstacles. One key barrier towards exploiting this source of energy in Uganda is the lack of accessible, authoritative and reliable geo-science data on geothermal prospects. Consequently, it is crucial that the technical personnel of GRD acquire the necessary training to improve their knowledge and skills in data management in order to competently carry out their mandate.

In collaboration with EAGER (East Africa Geothermal Energy Facility), the UK-funded program whose mandate is to facilitate the development of geothermal energy for power generation and other uses, the staff of GRD commenced the development of a relational database management system (RDBMS) that would efficiently store and retrieve geothermal data. One of the key benefits of the RDBMS is to enhance the GRD's capacity to plan, develop and manage concessions, with informed policies and regulations, which would in turn offer an environment conducive to outside investment in Uganda's geothermal sector. The GRD data management team is shown in Figures 1a through 1c.

## 2. Principles and Approach

A key aspect of geothermal development is properly gathering, managing and interpreting technical data (e.g., geographic, geologic, financial), as well as maintaining the data's integrity, in order that that it can be integrated and evaluated in order to reduce the technical and financial risk associated with drilling and well-field development activities. The best exploration risk mitigation is achieved by experienced data interpretation (IGA, 2013). In many cases, field intelligence and decision-making expertise from subject matter experts (SMEs, e.g., geologists, geophysicists, and reservoir engineers) is not sufficiently linked with government officials, policy makers, regulators and the many stakeholders. The planning and execution of successful commercial geothermal projects is critically dependent on reliable, trustworthy, and readily accessible data.



**Figure 1a (L).** Author geologist P. Maweje (2<sup>nd</sup> from right) with aerial drone for geophysics data collection. **Figure 1b (R):** Authors P. Maweje (far left) and geologist J. Nnakirijja (4<sup>th</sup> from left, seated) on Equator.



**Figure 1c.** Ugandan geothermal data team in field, with authors geochemist J. Achieng (far left) and geologist P. Maweje (4<sup>th</sup> from left). EAGER in-country support staff-geologist L. Nkunda-Ngabirano (far right).

## 2.1 Geothermal Data Types

Geothermal data can be highly diverse and wide-ranging, depending on the prospect and the level of exploration which the prospect has been subject to. This data can be structured, semi-structured or unstructured and may represent data from a fixed location or time or time-dependent data. In general, there are several broad domains of geothermal data, including:

1. Desktop studies (e.g., papers, reports, maps, either physical hard copies or soft copies)
2. Physiography and geomorphology
3. Geology
4. Geochemistry
5. Geophysics
6. Imaging and geodesy/geomatics
7. Wells and boreholes
8. Socioeconomic (e.g., environmental, ecological and cultural data)

These classes of information, while not comprehensive for any one particular geothermal prospect, are likely to characterize about 90% of the data that will be encountered over the life span of a project (Anderson, 2013). Some data tends to be unstructured, either not codified, or worse, not converted to machine-readable form. The key to successful development of geothermal power is to collect, manage and disseminate this data, based on best practices more widely than now, while simultaneously ensuring that confidential data obtained from developers is treated appropriately. The goal is to improve the capacity and decision-making capability of in-country stakeholders by providing them with access to data tailored to their specific interests.

The proprietary nature of much of the collected data is another challenge, due to (for example) the different formats and different emphases of different organizations, and the general non-transparency prevailing in the private sector, especially that dealing with the risky exploration for oil and gas (O&G) or geothermal resources. O&G is specifically mentioned here because the most relevant actionable intelligence to support a decision where to drill may not come directly from the geothermal investigations, but rather from a related field such as petroleum exploration. GRD takes this opportunity to acknowledge the valuable assistance it has received from its sister agencies. The “Blue Book” and “Yellow Book” for well and geophysical data standards and reporting used by the Petroleum Exploration and Production Department (PEPD) and Petroleum Authority of Uganda (PAU) became the model for GRD’s own “Green Book”.

*Illustrative example of the value of comprehensive and modern data management for decision support:* Many field reports up to a century old in cursive handwriting on miscellaneous paper exist in government/corporate archives or public libraries in East Africa. Despite archaic form and age, sometimes the content can be absolutely critical to a modern geothermal project. For example, a multi-million-dollar exploratory drilling project by a foreign aid donor would not have been located where it was had the sponsors seen the report of a foreign oil exploration project in the vicinity which had been undertaken over 50 years before. Interpreting the old drilling logs, and cognizance of a 600-meter offset between legacy coordinate systems (UTM based on colonial-era “Arc60”) and modern systems (WGS84), would have made all the difference. The relevant information existed in the form of a dusty report on a library shelf, and was known to exist by a few experts in the related O&G field. However, it was not widely known of, nor readily accessible to, the key decision-makers of the geothermal project.

## 2.2 Summary of Database Types

The types of databases (DB) that exist (in chronological order, from simple to complicated) are:

1. Flat files, which have existed since at least the 1800s.
2. Spreadsheets, consisting of simple row and columns (e.g., in the late 1800s in census records, then with embedded macros and functions when desktop automation began in the 1970s to present).
3. Hierarchical databases, consisting of “parent/child” tables, conceptually similar to e-mail folders. The multiple redundant entries required introduces the danger of inconsistency; also, the hierarchical architecture cannot support “many-to-many” relationships necessary for truly understanding and acting on one’s data.
4. Network databases (in which “nodes” mean collections of records, “relationships” equivalent to set structure, and the “owner/member” concept superseded the parent/child paradigm), followed hierarchical DBs. Network DB architecture supports “one-to-many” relationships, making them more powerful, but changing a relationship requires changing the set structure.
5. True relational databases, based on the mathematical principles of set theory and predicate calculus, arrived in the 1990s. Specialized terms include: “relations” which means tables, “tuples” which means records, “attributes” which means fields. The benefits of this approach are:
  - a. each record possesses a unique identifier
  - b. they provide the ability to manipulate and ask questions of the data (i.e., “query”) via structured query language (SQL)
  - c. with currently available software, queries can be generated intuitively in a GUI without needing programming skills to formulate the query.
6. After 2000, the “object-relational” model extended the RDB to object-oriented elements such as classes, encapsulation, and inheritance. Extensions allow an RDBMS to handle storage- and memory-hungry elements objects such as audio, video and architectural/CAD/GIS drawings and images.

## 2.3 Data Needs and Choice of RDBMS

As a matter of general principle and best practice, any successful information-processing system must be “needs driven” by the end users (bottom up), not by the architects of the system (top down). According to international best practices in data management, one of the first steps in designing a data system is to write a “requirements document” that specifies the overall purpose of the system (“goal”), then defines the data to be captured (“content”), the use to which it will be put (“functions”), and who will be using it (“users”). This requirement document leads to a conceptual model of the database (which should not be confused with the conceptual model of the geothermal resource). To be effective and “fit for purpose”, design of any RDBMS must be informed by “business rules” – the operating rules that govern a particular enterprise. Thus, a good RDBMS is not generic; it is highly context-specific. At minimum, business rules encompass the enterprise’s:

- mission, mandate, authority, and limits thereof;
- policies, such as those concerning administrative privileges of various classes of users, and confidentiality; and

- procedures, e.g., deadlines, reporting, chain of command.
- Example: In Uganda, the primary document instantiating a geothermal concession is the exploration license granted to a developer by MEMD. Therefore, the license DB becomes the highest-ranking DB. Geo-referenced boundaries and dates in the license DB thus become primary data elements in the schema, from which all other data elements flow, and to which all other data elements can be traced back.

The conceptual model in turn informs the decision to choose a relational or non-relational database. The advantages of an RDBMS over the other database types include:

1. Stored data is logically and physically independent from the application that create it or uses it. Neither changes to the DB design by users, nor changes by the vendor to DB application(s) will alter or harm the stored data itself.
2. Guaranteed data consistency and accuracy
3. Information retrieval as easy (or hard) as the designer wishes. The data can be viewed in nearly unlimited ways.
4. Data integrity built in on multiple levels, including:
  - a. at the field level to ensure consistency and accuracy,
  - b. at the table level to prevent duplication,
  - c. at the relationship level to assure validity between tables, and
  - d. at the enterprise level to assure data accuracy.

If various geothermal data constructs can be thought of as tools, then the RDBMS is the toolbox they are stored in. For these reasons, the use of a relational database makes the most sense in a geothermal context.

### ***2.3 RDBMS Design Phases***

The design of the RDBMS for the Ugandan GRD consisted of the following phases:

1. Requirements analysis, including:
  - a. Examination of the enterprise (GRD) being modeled
  - b. Interviews with management and SMEs to assess current and future needs
  - c. Assessment of information/data requirements for GRD
  - d. Skills assessment of enterprise personnel
2. Building the team, including:
  - a. Specification of team leaders
  - b. Assignment of Data Custodians (DCs)
  - c. Other key team assignments
  - d. Designation of Database Administrator (DBA)
3. Data modelling, including:
  - a. Entity-relationship diagramming
  - b. Schema definition
  - c. Data constraints
4. Normalization to collect, organize and then decompose large tables and data sets into smaller, more manageable ones to eliminate redundancy and duplication, including the cleaning of data before entry into the database.

### 3. GRD Database Development

#### 3.1 Implementation Plan

With the help of EAGER experts, an implementation plan was proposed and accepted for the development of an RDBMS for GRD that would include a front-end web-enabled interface coupled with the necessary skills development for GRD to build, populate and maintain the system through time. The implementation plan (see Figure 2) is a two-year program, with the first year focusing on skills development for the five key areas of RDBMS development:

1. Security
2. Operating system
3. Webservice
4. Database management
5. Web coding which includes both programming languages and content management systems (CMS)

The goal of this program is to develop the capacity within GRD for these data management activities. In addition, at the end of the first year, with EAGER’s assistance, the GRD shall have developed a limited but functional database with a web-enabled front end on the GRD website.

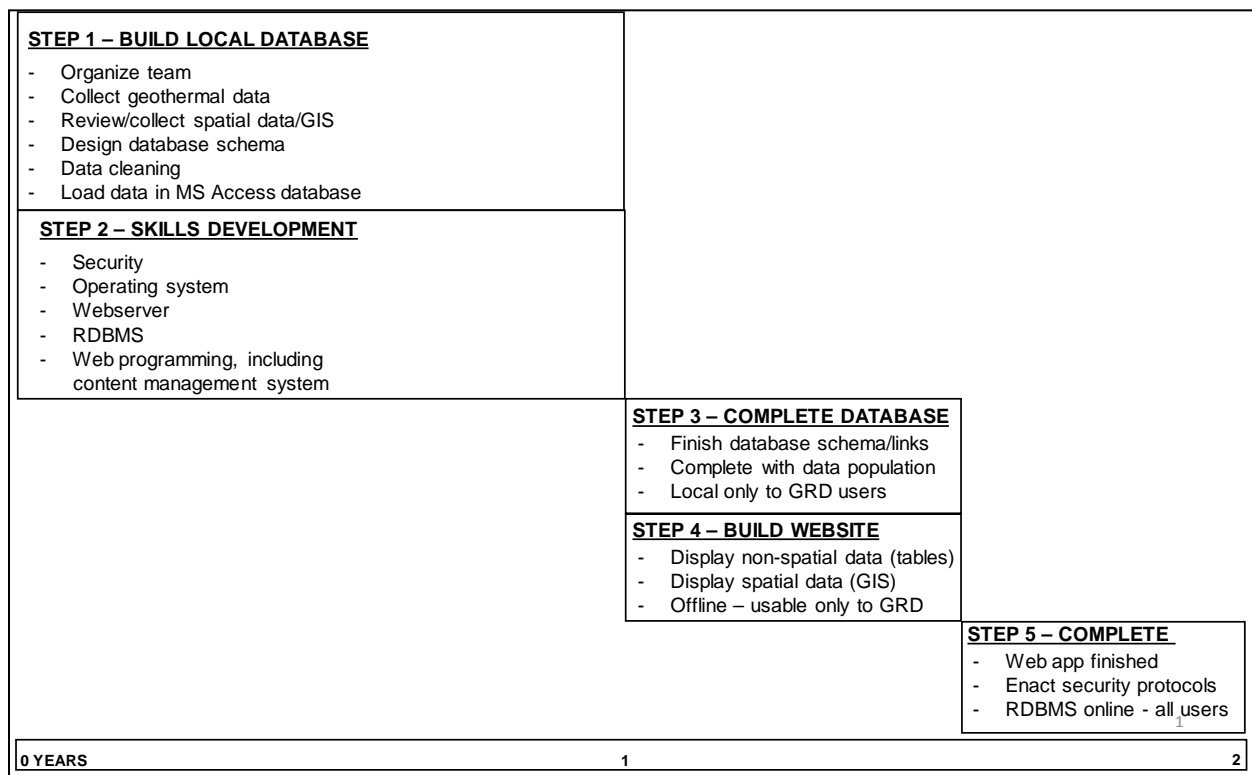
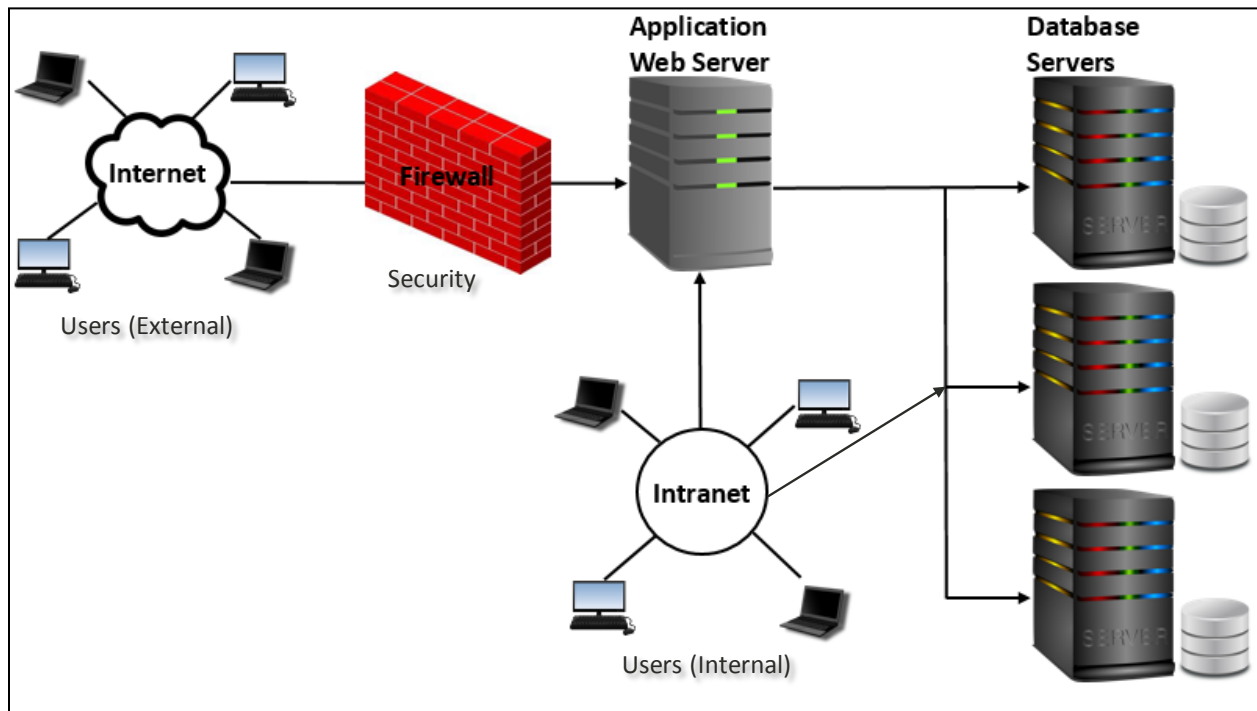


Figure 2. Implementation program for the development of the GRD RDBMS

These five elements, or layers, represent a technology stack (“tech stack”), that is, a development framework that includes all software subsystems to develop and run a full web-enabled platform for an enterprise (i.e., a full geothermal data management system), illustrated in Figure 3.



**Figure 3. Generic architecture for a database management system.**

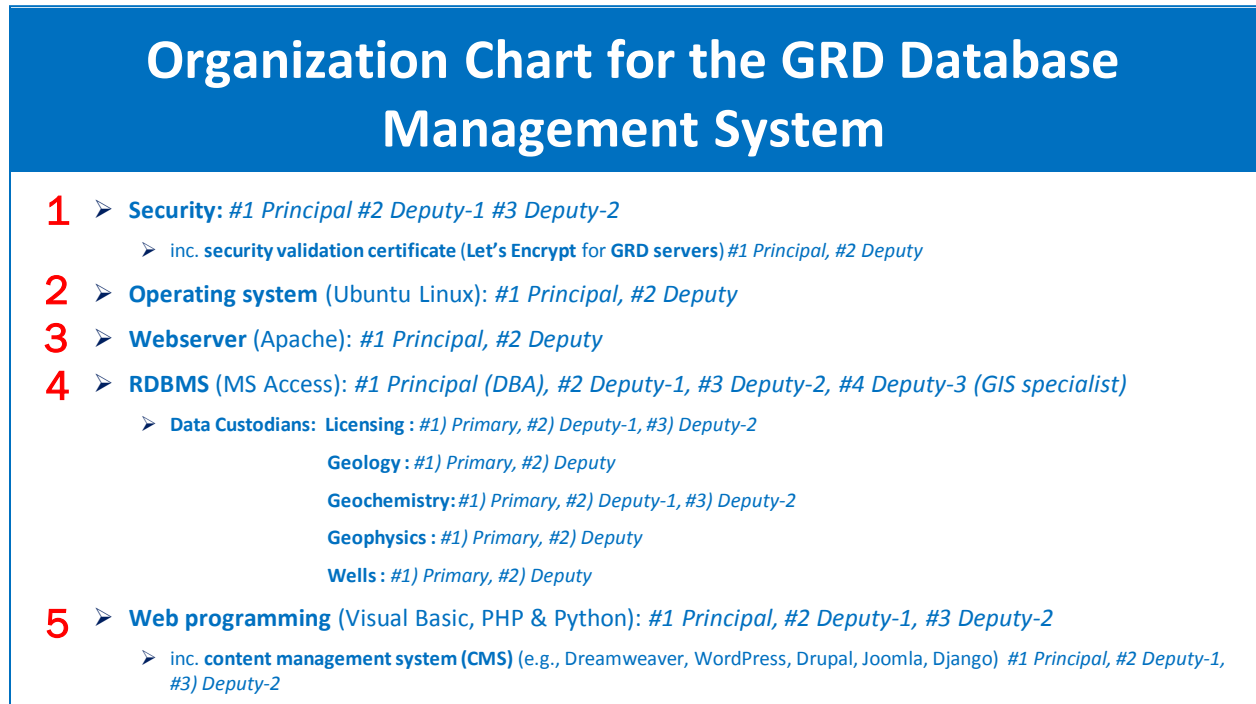
### 3.2 Data Custodians

A Data Custodian (DC) is the person responsible for identifying data sets within his/her area of expertise for storage in the database. This is the “point” person for that particular subject matter area (e.g., geochemistry, geology, geophysics). The DC is the person best suited to know and understand every form of data that exists for that particular domain (i.e., physical, electronic, paper, geo-referenced data, and indices). The DC is also responsible for the quality of the data (e.g., data cleaning) that ultimately will reside in the RDBMS. The maintenance, disposition and long-term preservation of records and data is another responsibility of the Data Custodian. To prevent single-point failure, it should be a two-person team consisting of a Primary and a Deputy. This provides continuity of operation and helps preserve institutional knowledge.

An organization chart was developed for each of the five layers of the “tech stack” as well as assigning DCs to each of the five main technical domains that comprise the RDBMS. In order of priority, these domains are:

1. Licensing
2. Geochemistry
3. Geology
4. Geophysics
5. Wells

The goal of team building was to provide redundancy at each layer of the tech stack, and domain, by designating a Principal and at least one Deputy. In some key areas, additional redundancy was provided by designating other GRD technical personnel as additional Deputies. The organization chart corresponding to enumerated skills in the “tech stack” is shown in Figure 4.



**Figure 4. GRD organization chart for the RDBMS development program.**

### 3.3 Data Cleaning

One of the key benefits of an RDBMS is assuring data integrity. This means the maintenance of the accuracy and consistency of data, primarily before it is imported permanently into an RDBMS. Data integrity ensures that all false information is excluded from the database while all true information is included in the database (Motro, 1989). The presence of incorrect or inconsistent data can distort the results of analyses, conceptual models, and resource models. One duty of the Data Custodian is to establish validity of the all the data before it is imported into the database, which is why the DC must be a trained scientist, not an ICT (information and communication technology) technician. Data cleaning is necessary early in the process to ensure data validity. Data cleaning (also sometimes referred to as data scrubbing) is the process of detecting and correcting corrupt or inaccurate records from a set of records or data tables. Errors to look for typically include (Hellerstein, 2008):

1. Data entry mistakes
2. Physically impossible values
3. Missing values
4. Outliers
5. Typographic and formatting errors (e.g., blanks, spaces)

After this process, the DC needs to ensure that the “cleaned” data is consistent with other similar data sets in the database.

### ***3.4 The GREDS Database***

The GRD data management team with assistance from the EAGER experts developed the schema for the geothermal RDBMS. The RDBMS was given the name “GREDS” (i.e., Geothermal Resources Department Database System). The GREDS database is a Microsoft Access database that contains a total of 31 data tables and 29 reference tables (i.e., look-up tables). Each table has a primary key defined, as well as foreign keys where appropriate. A primary goal of good database design is to remove data redundancy (i.e., duplicate data). To achieve this goal, data was divided into many different subject-based tables such that each piece of information was represented only once. This divided information can then be brought back together by placing common fields in tables that are related. These common fields in different tables are linked together by relationships.

In developing the structure of the database, it became clear to the team that the geothermal license granted to concessionaires by Uganda is the principal “information carrier” via which most other data elements are linked (e.g., geoscientific data), and thus necessarily should be the foundation of the schema. This is the basis of the GREDS database, and all data related to each particular geothermal license is linked together via the license identifier (i.e., geology, geochemistry, geophysics, and wells/borehole data).

### ***3.5 Linking the GREDS Database with GIS***

Much geothermal data is spatial in nature (e.g., faults, lineaments, hydrothermal manifestations, discrete sampling locations for geochemistry samples, etc.). Where appropriate, there is a unique numeric identifier for every record within each database that links data within and to other databases as well as to geospatial data in GIS. This obviates the need for data to be stored as duplicates in both the database and in geospatial software such as ArcGIS. Rather, geospatial data and their attributes can logically be stored and used in ArcGIS, but can be referenced via joins and relates back to their corresponding data elements via this unique identifier in the GREDS database. The identifier was given the working name GRID (GeogRaphic Identifier), until a name less prone to confusion is found. This data element is found in many of the GREDS data tables and has a corresponding data element (identical in value) in geodatabases for spatial data that is stored in ArcGIS. An example is shown in Figure 5.

The GREDS database is also being designed to access the UDIS (Unpublished Documents Information System) database within the MEMD. This database contains unpublished documents, reports and maps that have been scanned and indexed, and are accessible via SQL commands. In addition, the GREDS database will link to Uganda’s other legacy electronic databases (i.e., ERM and LIBERO). The information contained in these databases represents unstructured data (i.e., soft copy/scanned versions of documents, reports, maps, technical memorandum, etc.). As mentioned in Section 2.1, a well-designed database should be able to access multiple data types (e.g., unstructured, semi-structured and structured data). Providing the links from the GREDS database to these legacy document systems greatly improves the functionality of the database while eliminating the “disconnectedness” of the user (i.e., the

geoscientist) from accessing legacy information pertaining to a particular geothermal prospect of interest. Figure 6 shows the overall schematic of the GRD database management system.

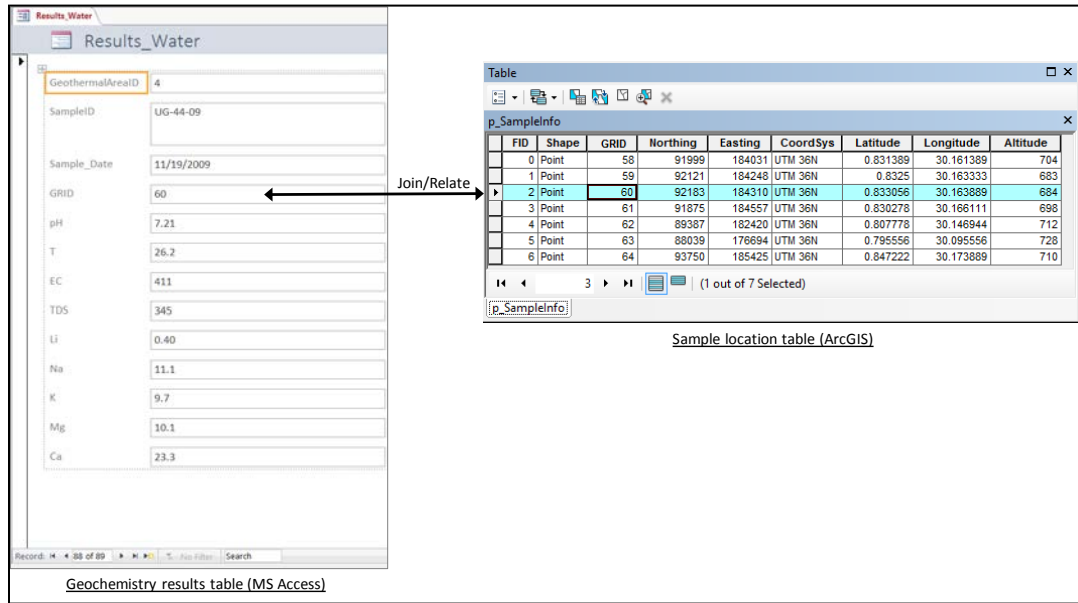


Figure 5. Linkage between RDBMS and geospatial data stored in GIS via the unique GRID identifier.

### 3.6 Linking the GREDS Database with Legacy Documents and Reports

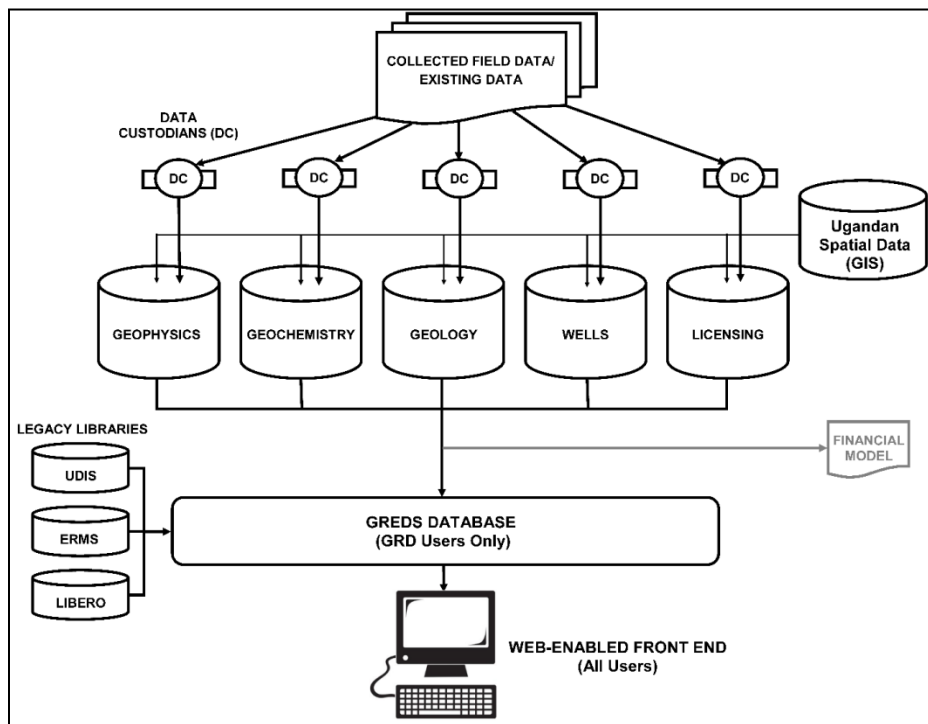


Figure 6. General schematic of the GRD database management system.

## 4. ICT Issues

The same phenomenon that restrains Uganda's overall economic growth hampers the GRD: the absence of reliable and stable electricity. The team has adopted a two-track approach to cope with this challenge, leveraging the deep understanding of each other's technical/managerial capabilities that has been gained over the past year-and-a-half. This in turn has driven immediate field-expedient workarounds and several major decisions regarding system design.

### *4.1 Lack of Reliable Electric Power*

Liaising with the Ministry's information/communication technology (ICT) personnel at the DGSM echelon via e-mail and Zoom video-conferences, EAGER is helping GRD to overcome or sidestep the issue of unreliable power supply. The team is developing specifications for a dedicated micro-grid to support critical components of the Directorate's ICT hardware systems. The team is also developing specifications for a system to respond instantaneously, e.g., batteries, in order to prevent data loss/corruption following unexpected but inevitable outages, which is the first priority; as well as another reserve system to operate for an extended duration system, e.g., generator(s), which is the second priority.

Liaising with GRD's geoscientists and DGSM's ICT personnel via e-mail and Zoom video-conferences, EAGER is also performing a cost-benefit analysis to assess the technical and economic feasibility of cloud-based or local storage (on DGSM's owned servers) of the geothermal RDBMS. The team is evaluating both internal factors and external factors. Internal factors include GRD's own capacity, cloud services offered by Uganda's National Information Technology Authority (NITA), and the Ugandan private sector's ability to provide required products. An external factor is the ability to reliably and cost-effectively access the worldwide cloud from defined locations in Uganda.

### *4.2 Virtualization*

One of the team's organizing principles is that it is better to get a limited but functional system up and running now, in order to start loading clean data and building competency right away, rather than to wait for a more complete, possibly better system later. Because the personnel environment can be dynamic, it is vital that all team members and DCs learn to validate and import clean data into the RDBMS as soon as possible. To facilitate immediate progress, two virtual machines (VM) have been instantiated in DGSM's server room, and set up under the geothermal.go.ug domain. Initially, they have been provided with 1 gigabyte (GB) of RAM and 50 GB of storage. They are named **files.geothermal.go.ug** (which hosts the raw data files, legacy data, files and reports) and **databases.geothermal.go.ug** (which hosts the GREDS database itself). The purpose of the files.geothermal.go.ug VM is for the Data Custodians to collect, organize and store all relevant but "siloed" geoscientific data still residing on multiple laptops, computers, and external hard drives onto this one machine. This includes existing and legacy data (e.g., Excel files, raw and processed data), GIS files (i.e., shapefiles), maps, figures and reports, stored in a consistent and logical file structure with standard file naming conventions. The first VM facilitate the cleaning of the data by the DCs before importing that data into GREDS, which resides on the second separate VM.

In addition, DGSM provided SSH access to the virtual machines to the EAGER. Secure Shell (SSH) is a cryptographic network protocol for operating network services securely over an

unsecured network. The best known example application is for remote login to computer systems by users. SSH provides a secure channel over an unsecured or secured network in a client-server architecture, connecting an SSH client application with an SSH server. Common applications include remote command-line login and remote command execution, but any network service can be secured with SSH. This feature allows the EAGER team to remotely monitor progress of GRD staff as they use the VMs for data storage and to build/populate the DB. This architecture, diagrammed in Figure 7 to show VMs in relation to the world, provides:

- secure data retrieval and access control,
- maintenance, disposition and long-term preservation of records and data and
- improved quality control.

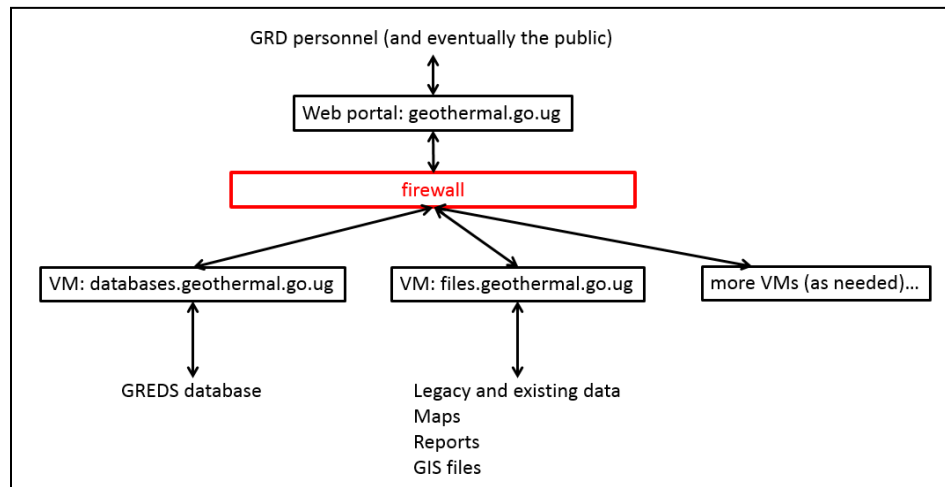


Figure 7. Block diagram of GRD's domains and virtual machines.

These VMs were then mapped to be visible on DGSM's local area network (LAN) from the desktops of the project leads at GRD: P. Maweje, J. Achieng, and J. Nnakirijja. See Figure 8.

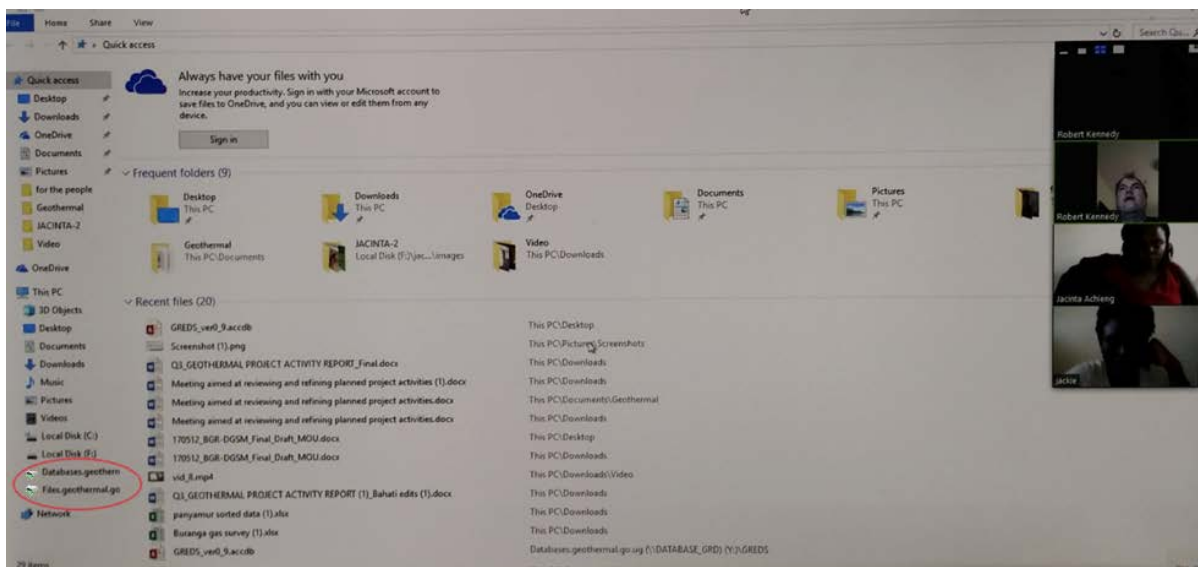


Figure 8. Connectivity established to virtualized servers comprising the RDBMS, as it appears on a GRD desktop computer during international Zoom teleconference across 8 time zones.

Figures 9a, 9b, and 10 illustrate the process of setting up these VMs.

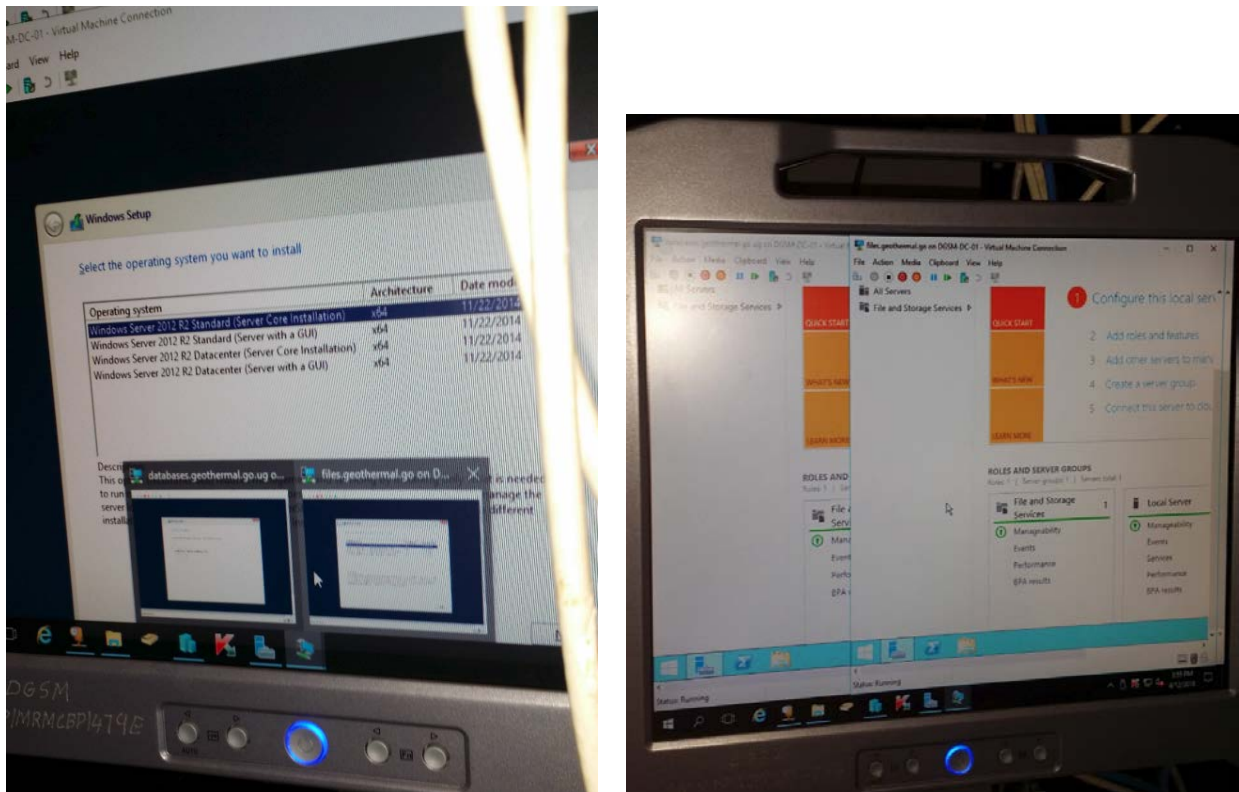


Figure 9a (R) and 9b (L). Setting up virtual machines for hosting data files and GREDS database.



Figure 10. Setup of virtual machines complete. From L to R: M. Nganda (DGSM ICT lead), D. Mulinde (PAU), author-geologist J. Nnakirijja (GRD DBA), and author-hydrogeologist J. Benegar (EAGER).

Figures 11-13 illustrate security certificate installation and the resulting secure GRD website.

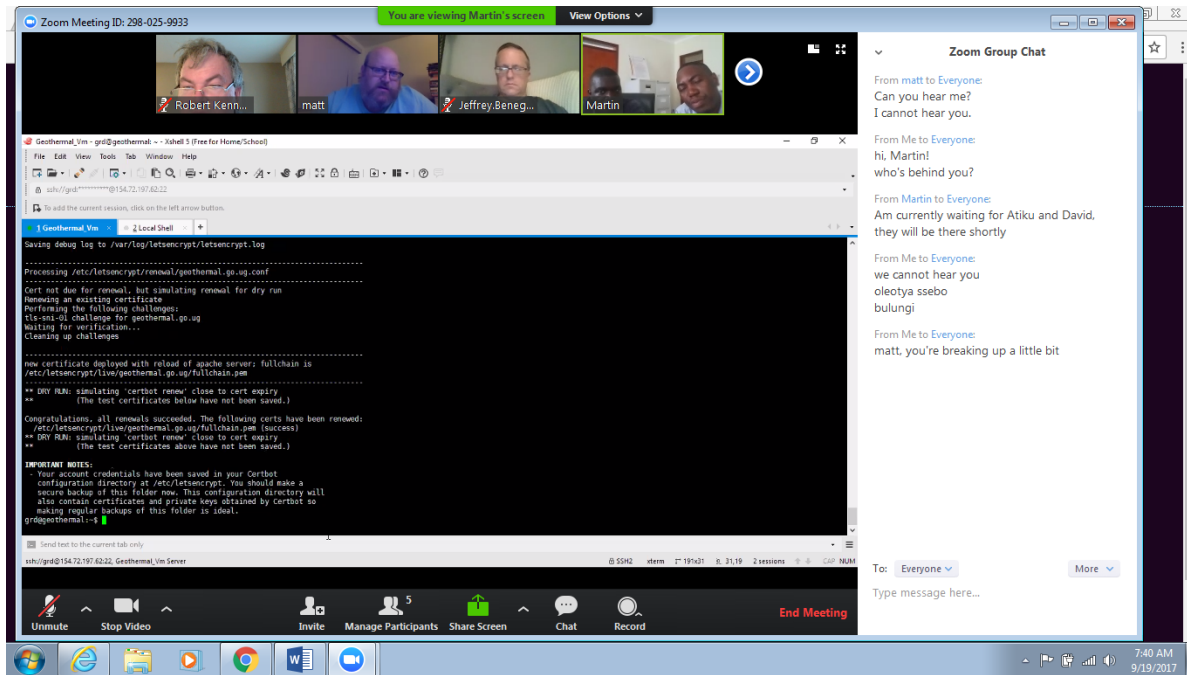


Figure 11. Installing “Let’s Encrypt!” security certificate on GRD’s webserver for domain geothermal.go.ug via teleconference across 11 time zones. From L to R: author R. Kennedy (EAGER), M. Johnson (security consultant), author J. Benegar (EAGER), E. Atiku (consultant) and M. Nganda (DGSM ICT)

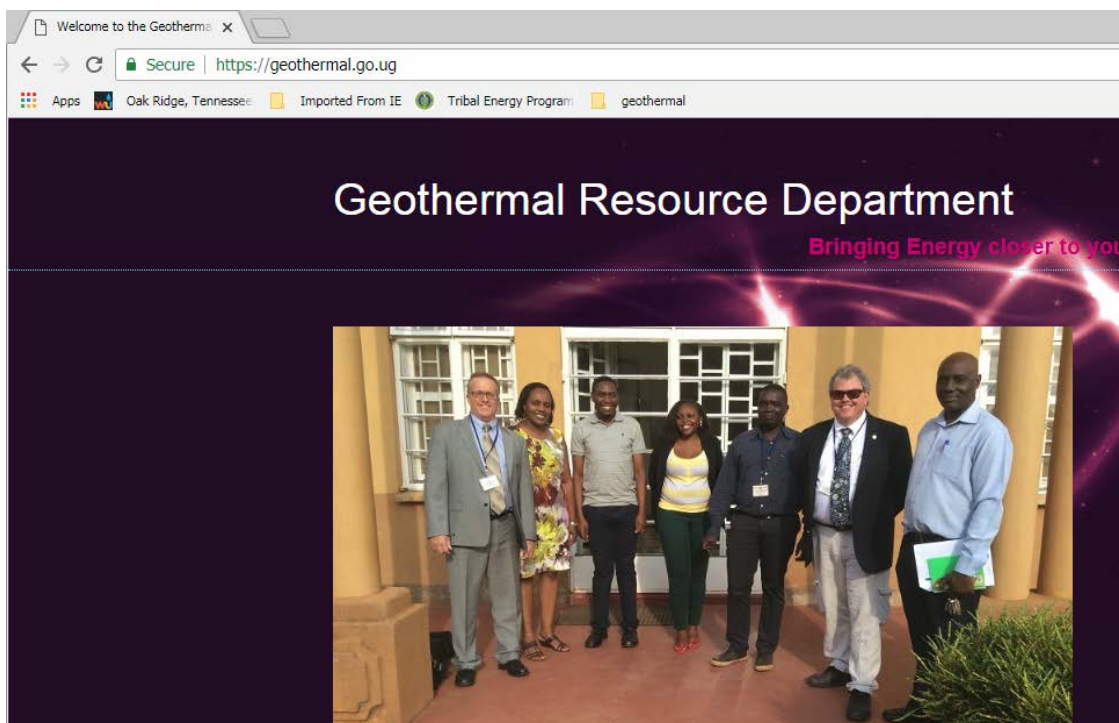


Figure 12: Author J. Benegar (EAGER), L. Nkunda-Ngabirano (EAGER). M. Nganda (DGSM), author J. Achieng (GRD), E. Atiku, author systems engineer R.Kennedy (EAGER), Act.Comm. G. Bahati (GRD)

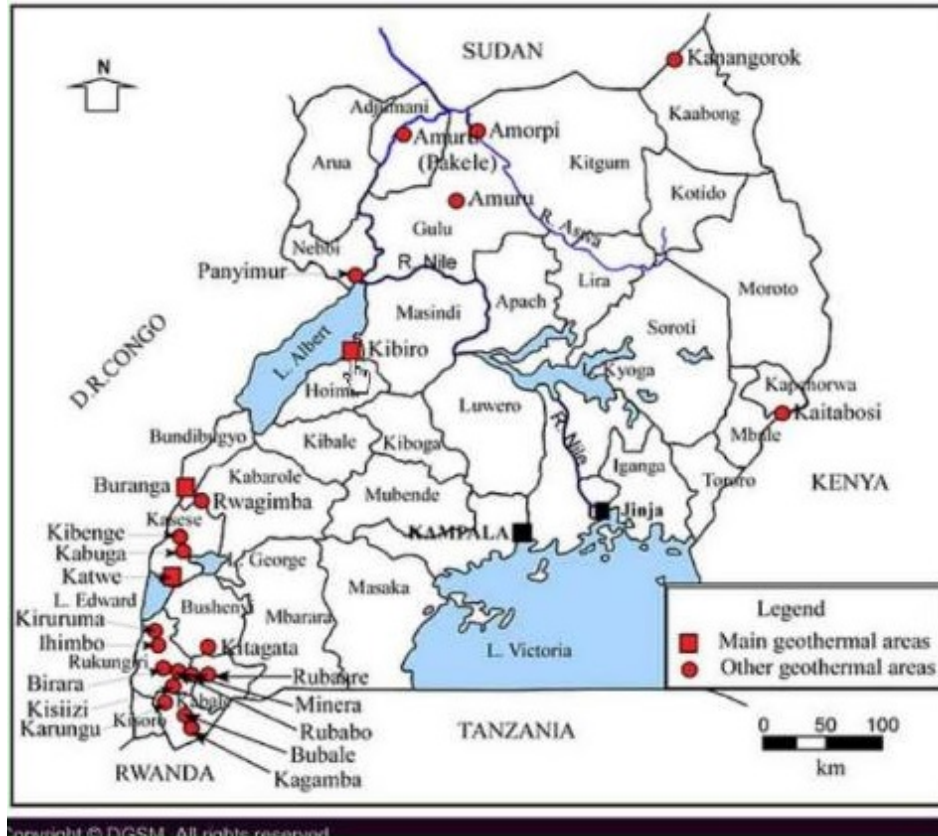


Figure 13. Current GRD webpage with added geographic map element of selectable geothermal concessions.

## 5. Conclusions

Uganda's Geothermal Resource Department, with assistance from the "EAGER" program of the UK's Department for International Development, has implemented a limited-yet-functional, secure, and modern relational database management system. A comprehensive schema based on similar exploration activity taking place in the oil and gas sector, encompassing numerous disparate and interrelated data types, has been laid out. Virtualized servers, custom domains, and a secure website have been set up to host the RDBMS. The RDBMS enable users with different devices and privileges to appropriately and securely access geothermal data from a variety of sources, e.g., field work, geospatial information, and legacy reports concerning the different geothermal prospects of Uganda. A great deal of legacy data, formerly "siloes" among individuals and departments, is being gathered together, cleaned, validated, and backed up. A robust team of trained geoscientists, with enough redundancy to prevent single-point human resource failures, has been organized to operate this system. Existing scientific, data management, and ICT skills have been broadened and strengthened across a very broad spectrum of disciplines, e.g., licensing, geology, geochemistry, geophysics, reservoir and drilling engineering, and GIS. Substantial new technical skills have been acquired as well. At the same time, these skills, old and new, are also being copied and shared via cross-training ("all-rounders") to provide further resiliency to the Department, the Directorate above it, and

ultimately to the country of Uganda, in order to manage the nation's geothermal concessions and correctly inform important decisions, such as where to drill.

## REFERENCES

- Anderson, Errol. "Developing a Universal Geothermal Data Management System." GRC Transactions, Vol. 37 (2013).
- Hellerstein, J.M. "Quantitative Data Cleaning for Large Databases." United Nations Economic Commission for Europe (UNECE) (February 27, 2008).
- IGA. "Geothermal Exploration Best Practices: A Guide to Resource Data Collection, Analysis, and Presentation for Geothermal Projects." IGA Service GmbH and IFC (2013)
- Motro, Amihai. "Integrity = validity + completeness" ACM Transactions on Database Systems (TODS), Volume 14, Issue 4 (December 1989).
- National Planning Authority. "Uganda Vision 2040", <http://npa.ug/wp-content/themes/npatheme/documents/vision2040.pdf> (2007).