# Geotechnical Data Management using acQuire Workflows and Data Integrity

M.J. Murphy
*Golder Associates Ltd., Vancouver*

B.A. Murphy
*SRK Consulting, Vancouver*

ABSTRACT: acQuire is a database management system for the storage of quality data related to the exploration and mining industry. Poor data management systems can lead to inconsistencies and loss of data which results in a waste of time and money to achieve a usable set of data. Integral to the acQuire system is the data model with its different components which allows for the storage of data collected at the different scenarios in the exploration program. acQuire is commonly used for organizing geological and geochemical data but is now being used for the management of geotechnical data sets.

This paper will discuss possible work flow models. Also discussed will be quality control and data validation aspects from the geotechnical logging and mapping phase through to the integration of the data in the database and the running of procedures to perform calculations; validating the data and the retrieval of the data for reporting and other geotechnical procedures.

## 1 INTRODUCTION

Paper logging systems have long served a purpose but have been replaced over the years with more modern digital systems following the trend and development in Information Technology and the associated software industry. In situations where a paper system is still being used, the data collected will inevitably be transferred to a digital scheme for storage and use in application products. It therefore makes sense to capture the data digitally at the onset of a project, making use of features such as pick lists with valid selections and comment fields for extra information. Digital data capture systems encourage good data collection protocols by establishing business rules and validation procedures.

Personal Digital Assistants (PDA) are handheld devices originally designed as personal data organisers but have developed into devices with data capture functionality. While these have a definite use for exploration geologists who cover a vast area and for whom portability is critical, they are limited in the amount of data that is visible on the small screen. These limitations exist not only in the number of fields and rows that can be displayed but also in the number of characters and are therefore more suitable for collecting coded data. An additional limitation is with the data collector losing touch with what they are collecting. If portability is not necessary, a bigger screen simplifies the data collection process immensely and is therefore preferable.

When capturing geotechnical data in the core shed or even at the drill rig, a laptop computer provides a user-friendly solution to capturing digital data. The decision to use such a device requires a re-assessment of logging methods and the introduction of concepts of human-computer interaction. Ideally the interface should be intuitive for all users, allowing any input while maintaining data integrity and providing a method for real time feedback to the logger to help identify errors. Some data capture interfaces are analogous to 'black boxes' where the logger enters the data but does not see a final log and loses the control that was had when capturing the data on paper. In addition, a digital interface can present QC graphs to the logger enabling them to recognise and hopefully rectify errors in their data. "Digital data handling forces us to adopt strict rules on the uniformity of data formats which is essential for the "interoperablilty" of data on different parameters and from different data sources." (Rengers et al, 2002)

# 2 DIGITAL DATA CAPTURE

## 2.1 *Spreadsheets*

Spreadsheet applications (or 'spreadsheets') are computer programs that allow the creation and manipulation of tables of numbers and/or text, with individual values stored in a 'cell'. Cells can be linked together by means of formulas which will come into effect as values are entered into the cells.

Commonly known spreadsheets such as Microsoft Excel or Lotus1-2-3, are often used to capture data digitally. These systems are advantageous in that a large number of people have access to the programs which are functional, relatively cheap and also fairly well known to the vast majority of people. "Decision makers need to analyze data, and they know excel." (Chung, 2004) Some validation can be written into these spreadsheets and individual boreholes can be logged quite soundly using a spreadsheet application. As the amount of data increases, however, software limitations result in the data problems discussed below.

## 2.2 *2.1 Classic issues with spreadsheets*

Very few people can say that they have never struggled with their data. For the most part, those using raw data to run calculations and make decisions will certainly have experienced various problems, one of the most likely is the existence of many data files and the uncertainty as to which is correct and most recent. Other commonly encountered problems relate to a lack of confidence in the data result in the following questions:

- How do I know that there are no repeating or overlapping intervals?
- Do I have any anomalous intervals (FROM > TO)?
- Do I have data integrity? (An example of poor integrity would be where a drillhole name is stored differently in two different sheets.)
- Are the same codes being used throughout for data fields such as lithology? (E.g. G, GR, granite and Granite might all mean the same thing but will not be returned in a search.)
- Is it possible for more than one person to access the data simultaneously?
- Can I delete or edit a drillhole name without having to refer to numerous sheets? Similarly with editing or deleting intervals.
- How easy is it to combine data sets captured at different intervals? (E.g. run data and domain data captured during geotechnical logging.)

Finally there is a limit to the number of records that can be stored in a spreadsheet. An individual drillhole may be logged in a spreadsheet but will be added to the other drillholes in the project. To accommodate the growing amount of data and to avoid the uncertainty presented with a multitude of spreadsheet files, a relational database would be the tool to use when spreadsheets become unmanageable.

## 2.3 *Relational databases*

A database where the data is stored based on its relationship to other data in the same database is termed a 'relational database'. Data needs only be entered once and can be retrieved in any number of formats. Deleting or editing information like a drillhole name need only be done once. Microsoft Access is a relational database and its use eliminates the problems associated with spreadsheets as discussed above.

To appreciate why a relational database is important, it helps to look at drillhole data, which for many conjures the image of a single large table of fields and records displaying much information about the drillholes. This is a superb way of viewing the data but not an efficient method of storing and managing it as inherent weaknesses are introduced. Some weaknesses in storing data in this manner are:

- Duplicate data can occur.
- Repetition of certain data is introduced which is a waste of storage space. (E.g. Collar data may be repeated for multiple records.)
- Empty cells exist where there is no value for a particular field. This results in wastage of storage space.
- There is no cascade delete or edit. (E.g. a drillhole deleted in one sheet might still occur in another.)
- Filtering and presenting the data is limited. (E.g. viewing data from specific holes or joining data collected over different intervals.)

A relational database overcomes these issues by introducing concepts such as tables and fields; relationships amongst the data fields; primary keys and normalised tables.

## 2.4 *2.3 Database Design*

A desktop user can make use of Microsoft Access to efficiently create a powerful relational database application for manipulating data. A relational database can be designed with the goal of collecting geotechnical or other geoscientific data, and data input forms can be generated for collecting the data. Access provides user-friendly wizards that can assist in getting started or in searching for data in an existing database.

of the project's data then one needs to look for something more powerful than Microsoft Access.

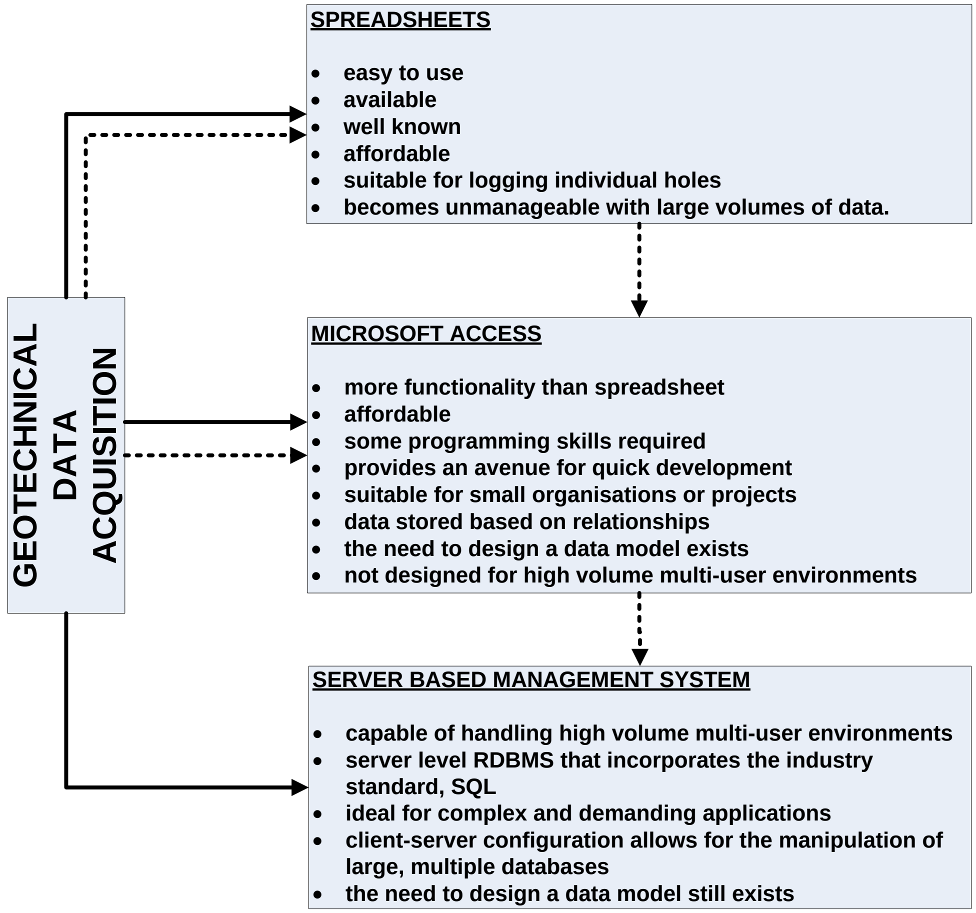


Figure 1: The Advantages and Shortcomings of Different Digital Capture Systems

For small projects with limited budget, Microsoft Access can provide a technically sound solution to the problems associated with managing and storing geotechnical data. "If there is a lot of data, SQL Server is the better choice. SQL Server also supports more users and traffic." (Chung, 2004) SQL server is also ideal for combining several Access databases.Figure 1 shows the different data transition paths that could be followed. A project might start with Excel and migrate to Access and migrate further to a server based system or else it could utilize a server system from the onset. As stated by Luke Chung: "When an Access application is created initially, the features needed in the future cannot be anticipated." Optimal database size with Microsoft Access is potentially 750MB or 500 000 records. As the database grows in size, it becomes inefficient, especially when performing calculations which can be time consuming.

If geotechnical data is the only type of data to be managed then the structure of the database need not be overly complex but just address necessary fields for collar and geotechnical data. If a database is likely to be large (greater than 750MB) and have the need for multiple simultaneous users, as well as be able to accommodate different aspects Microsoft's SQL Server or Oracle are both server-based relational database systems that can handle large amounts of complex data and will therefore address the issues of size and number of users. The need to incorporate data such as samples despatched to a lab; assay results; downhole or surface geology and survey data will require major changes to the existing structure of the database.

"If you decide to store the data for your solution in a database or in another relational database, designing the database structure (DATA MODEL) is likely to be the most challenging part of building the solution." (Microsoft web site).

This therefore prompts the question: *Is there a commercially available product developed for the capture, management and storage of quality geoscientific data in a database system that can cope with multiple simultaneous users and handle a large volume of data as well as having the capability to store other types of data?*

The acQuire data management system is such a product which strives to eliminate the problems encountered in the capture and management of geoscientific data and thus merits closer scrutiny.

# 3 THE ACQUIRE DATA MANAGEMENT SYSTEM

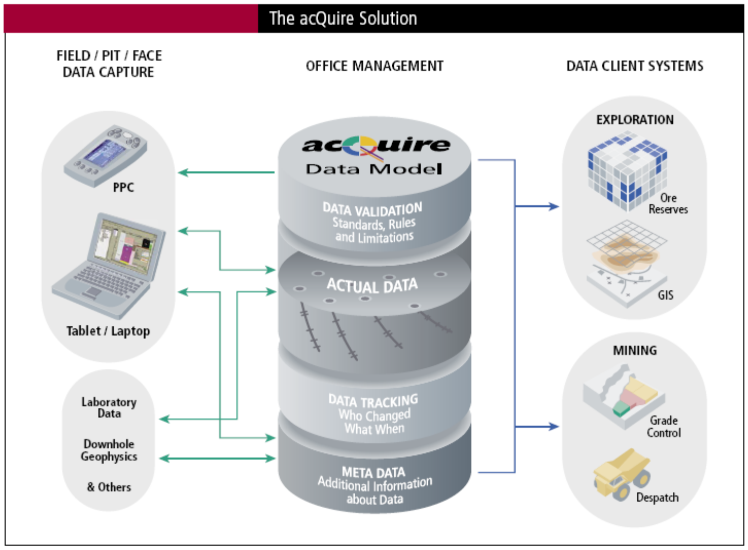


Figure 2: Taken from acQuire brochures showing acQuire technology and illustrating the link between the data capture systems, the acQuire data model and the application systems that utilise the data.

## 3.1 *3.1 acQuire data model (ADM)*

Central to the acQuire package is the acQuire data model (ADM) which is designed to operate in the SQL and Oracle environments. As alluded to earlier, designing a database structure (or model) is likely to be quite challenging, time consuming and possibly too specific for future enhancements. The acQuire data model is a set of database tables and relationships which relate to each other in the best possible way, adhering to data management best practices as well as what is logical for retrieving and managing exploration and mining data. One of the many advantages of using a standard model is the ease with which data can be transferred among different sites as well as reducing the need for training and reliance on key personnel. Additionally, a project may start using one area of the database, such as surface sampling, then move to assay results and despatch data and then move to drillhole geology, geotechnical and structural data. The acQuire data model is designed to incorporate all of this.

A schematic diagram of the acQuire data model is provided with the software and illustrates the tables that exist in the different sections of the model. Inherent in the design are features such as a data loading order; cascade update and delete; primary keys and the relationships between fields. An understanding of the model is necessary to work with the system and training to managers and users is available.

The data loading order enforces data integrity, e.g. data cannot be entered into an interval that has not been defined nor into a drillhole that likewise has not yet been defined. The cascade update and delete function assures that if data (e.g. HoleID) is edited, or deleted, where it was defined then the change will be reflected in all tables where that data appears without the user having to search for all occurrences.

## 3.2 *3.2 acQuire front end functionality*

In addition to the ADM, acQuire has a front end with a suite of tools used for inserting, manipulating, retrieving and presenting data. Each project is different and will make use of different tools while customizing them for the individual needs. New products are added with new versions of the software and existing functionality may be enhanced. These tools can assist in establishing a site-specific workflow.

acQuire licensing enables users to perform different tasks, e.g. a manager licence will allow site-specific fields to be created and validated in the database. Likewise, a data entry license will allow the remote capture of data into a ‘stub’ of the database which can be imported into the main database with little effort. An acQuire security device is necessary to use the software unless a network license is owned by the company.

# 4 USING ACQUIRE TO COLLECT GEOTECHNICAL DATA

Although acQuire is an ‘out of the box’ system, the implementation thereof requires careful planning involving data managers, geologists and engineers. AcQuire Technology Solutions (ATS), the vendors of the acQuire product spend quite a bit of time in the sales process scoping the solution with prospective customers in order to provide adequate costing as well as scheduling the implementation.. A subset of existing data is also examined in order to assess how long it will take to migrate it to acQuire if needed. Questions to be considered are:

- How many users are there and what will they be doing (data collection, data analysis etc)
- What is the state of the current data, and how much of it needs migrating to acQuire?
- What are the current workflows and what needs to be set up to emulate this?
- Where is the site located and what are the communications between sites (if multiple sites) like?

The answers will determine the time, effort and cost involved.

There are some major mining companies, such as Anglo American and Rio Tinto who have standardised on the use of acQuire at exploration and mine sites. This means that similar collection interfaces exist and personnel have applicable knowledge when arriving at a new site. Data can be transferred from site to a central or 'main database' at head office or at exploration offices. As with all systems, an acQuire system needs to be managed and some level of acQuire expertise is required.

If an acQuire system has already been established and one wishes to add the geotechnical component, much of the 'ground work' in terms of business procedures and workflows has already been completed. Geotechnical parameters can be added to the database and collection interfaces altered. Once the data is entering the database it can be viewed or extracted with any of the existing data.

## 4.1 *Logging and Mapping*

acQuire has the capability to create a data collection interface which can be used remotely on a PDA or a laptop. The interface is exported from the main database, ensuring that the business rules built into the database are applied at the remote site where logging is taking place. Data collection for the loggers is simplified with the use of drop-down lists providing them with valid choices for the particular field. It must be noted that the logging interface creates a temporary access database on the remote machine and installing SQL server is not required.

When exporting an interface from the main database it is possible to export a blank interface or to include existing drillholes for the reference of the logger. It is also possible to export a subset of validation fields which is useful when data is specific to a particular project. Project-specific business rules can be built into the interface initiating the first phase of quality control.

The iterative process of logging with the acQuire interface is shown in figure 3. Saving the file to an external device such as a memory stick or a portable hard drive provides a backup in case the working file is lost.

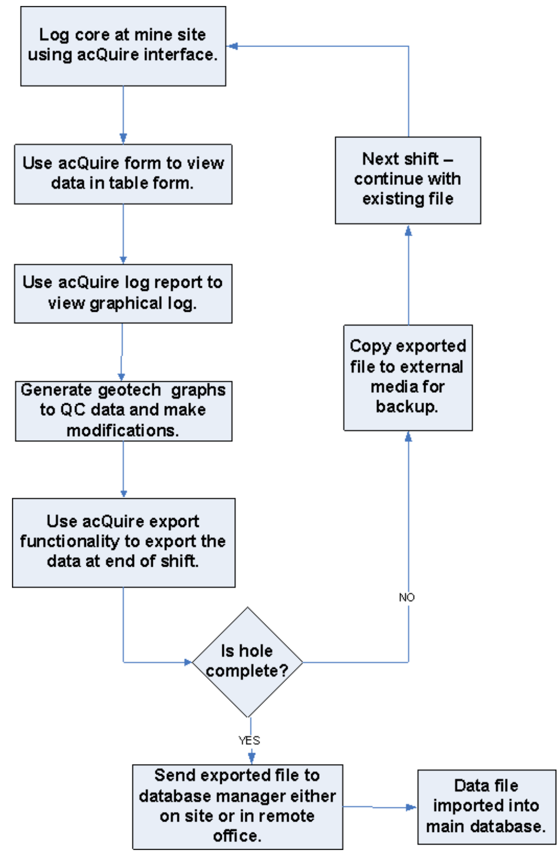

Figure 3: Proposed data collection workflow using the acQuire system.

If the logger is expected to perform rudimentary checks on the data collected it is beneficial if they can visualise what has been logged, either in a table form or else in a graphical log. acQuire has a log report object that can be created in the main database and exported with the interface for use by the logger as well as a form object which displays the data in a table format.

## 4.2 *Geotechnical Parameters*

### 4.2.1 *Drill hole Data*

Aspects such as the site name and project number; collar location of the drill hole; azimuth and dip of drill hole; the dates of start and finish of the drilling and the drilling company are collected once for each drillhole. The size of drill core; the type of drilling system used as well as the types of drilling-muds used might also be captured. When logging is undertaken on the drill rig, a performance sheet may be kept which indicates the total drill time, stand-by time and depth.

Incorporating this data in the acQuire system, the HoleID and any information about each individual drillhole would be captured once and stored in the Collar area of the model. A hole is uniquely defined by the combination of HoleID and Project Code which allows for the same HoleID to be used in different projects in the database.

#### 4.2.2 *General Geotechnical Parameters*

This stage of data collection involves the gathering of a base set of geotechnical parameters, captured with minimum interpretation. Parameters collected on a drilling run basis include logging intervals, recovery parameters and Rock Quality Designation (RQD) in the case of rock. If the project were at an early exploration phase, other aspects such as Intact Rock Strength (IRS), total fracture and joint count would be included for rock core; and type, strength, consistency and colour for weak rock or soil.

Data collected over an interval will be initially defined in the GeoInterval table with the value stored in GeoDetails. One must note, however, that the tables and relationships are the domain of the data administrator and are hidden to the user collecting the data.

When defining specific fields the database administrator needs answers to questions such as: "Will this data be captured at an interval or at a single depth?" and "Can I allow overlapping intervals for this data?" as different permutations can be enforced.

#### 4.2.3 *Detailed Geotechnical Parameters*

As the project advances and the level of study increases, the required level of data collection must also increase. The initial level of general geotechnical parameters will be insufficient to implement rock mass classification systems and undertake the more detailed geotechnical studies (e.g. pit slopes and underground excavations). For these studies, additional data on the level of micro-defects; fill strengths; number of joint sets, their orientation and the joint conditions in the case of rock and additional parameters for soil logging are required.

This increased level of data acquisition is easily accommodated with the addition of fields to the geology area of the acQuire data model.

#### 4.2.4 *Down-Hole Geophysical Data Sets*

An additional data set that can be acquired during the advanced exploration phase is the dataset from the downhole acoustic and optical TeleViewer as well as that from a full-wave form sonic probe. The geological features and discontinuity data sets, borehole break-outs, as well the various determined strength parameters will all need to be stored separately and then later merged with the geotechnical logging data-set to establish a representative geotechnical data set to be used in the design process.

Again, this data acquisition can be accommodated within the existing data model.

#### 4.2.5 *Field and Laboratory testing*

Accompanying the increased level of geotechnical data is the collection of field- and laboratory-based test results which include point-load tests, density, and weathering tests as well as the plethora of soil and rock geomechanical tests. The results of these tests need to be stored at the appropriate interval in the database and in acQuire this information would again be stored in the GeoInterval and GeoDetails tables.

Although it is preferable at the onset of a project to know all the data fields that are to be captured, the design of the ADM and the tool for creating fields is such that adding new fields to the database at any stage is fairly straightforward for someone with some basic acQuire training. The same logging interface can be used but enhanced with the new data fields.

#### 4.2.6 *Calculations*

Included in the acQuire tools is the concept of derived and calculated fields which can be used for determining any of the derived parameters from the initial set of basic geotechnical data. This can be designed to include conversions of parameters as well all the elements of the various rock and soil classification systems. Calculated fields are static and store the result of a calculation, whereas derived fields run the calculation each time the particular field is referenced, thereby optimising storage.

## 5 QA/QC PROCESSES

### 5.1 *Field Program*

It is essential that the major part of the QA/QC function should be undertaken during the initial geotechnical logging. This can be achieved with thorough conventional methods of training; a comprehensive geotechnical logging atlas including pictorial standardised logging of representative core boxes; and by using a very tight set of pick lists for each parameter to be measured. It is essential that the initial data-set is of a very high quality before it is passed on for post processing by personnel who have probably not been to the site or have no experience logging core from the site.

What has been noted on a number of geotechnical programs at various sites is that the geotechnical loggers do not have a feel for the data sets as a whole. It is often a case of entering a selection on a pick list with a view of a limited row of reference information. In other cases the geotechnical information is loaded daily into a database and is either used or stored off site. This results in the logger having a very limited view of the various parameter ranges of the drillhole as a whole or even within the same interval.

A potential method of overcoming this is to provide on site QA/QC by including a number of graphs that could be reviewed at regular intervals.

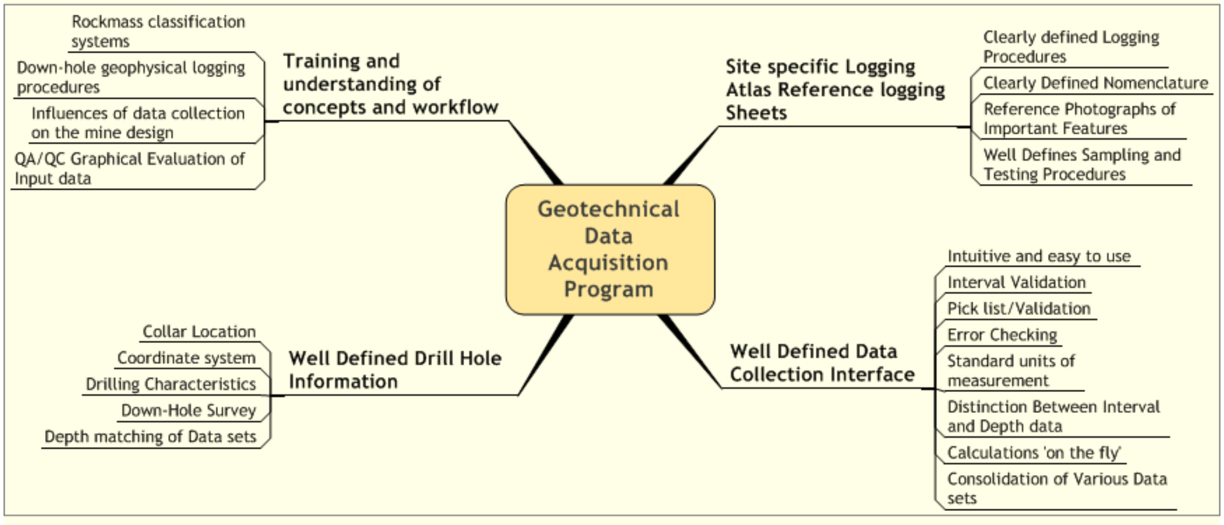


Figure 4: Elements of Good Data QA/QC on site

The following graphs are recommended for inclusion in the site logging forms:

- RQD and Total Core Recovery (TCR) percentage against the down-hole length;
- Intact Rock Strength, alteration intensity and rock type against down-hole length;
- RQD percentage and FF/m against down-hole length.

Regularly reviewing these graphs will highlight inconsistencies within the data set as well as allowing corrections on-site where, in the worst case, the core can be re-reviewed. acQuire allows the logger to edit data within the same interface that the data was initially collected.

### 5.2 *QA/QC Factors in the main Database*

#### 5.2.1 *Correlation of datasets*

Data gathered from the down-hole logging tools needs to be merged with the data-set from the detailed geotechnical logging. Business rules need to be established for this merging process, for example where depths of features differ, then the down-hole logging depths take precedence over those captured in the geotechnical log. In terms of features, however, that captured by the geotechnical logger takes precedence.

Another example of the requirement for business rules is in the case of intact rock strengths, where values are determined empirically during logging, from point load tests, from down-hole logging tools and laboratory tests. Establishing these rules may become complicated and may require the intervention of experienced geotechnical personnel.

It is essential to maintain the integrity of the original data and this is achieved in an acQuire

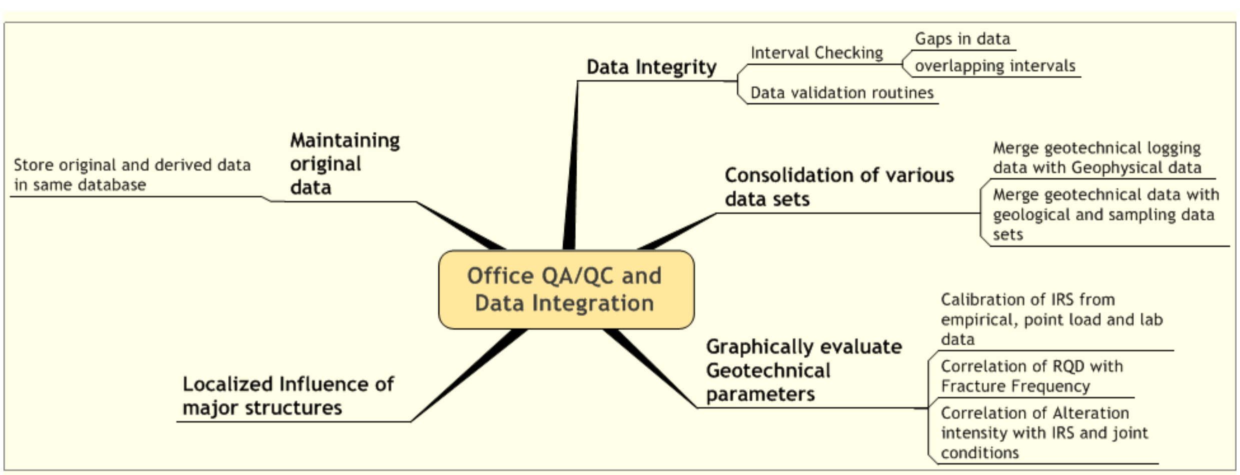


Figure 5: Suggested QA/QC Practices Prior to using the dataset for engineering design.

database by storing both collected and altered data.

### 5.2.2 *Data integrity testing*

Values may have been edited according to the business rules and thus the graphs generated on site should be re-generated and examined. In addition, another QC graphs to be viewed at this stage include IRS, Joint conditions and alteration intensity.

maintained and developed on an ongoing basis by a company rather than a specific individual.

The acQuire system does not currently fulfill all the requirements for geotechnical data management but through ongoing development of the application and continued interaction with geotechnical personnel this will become a system of choice for geotechnical data management.

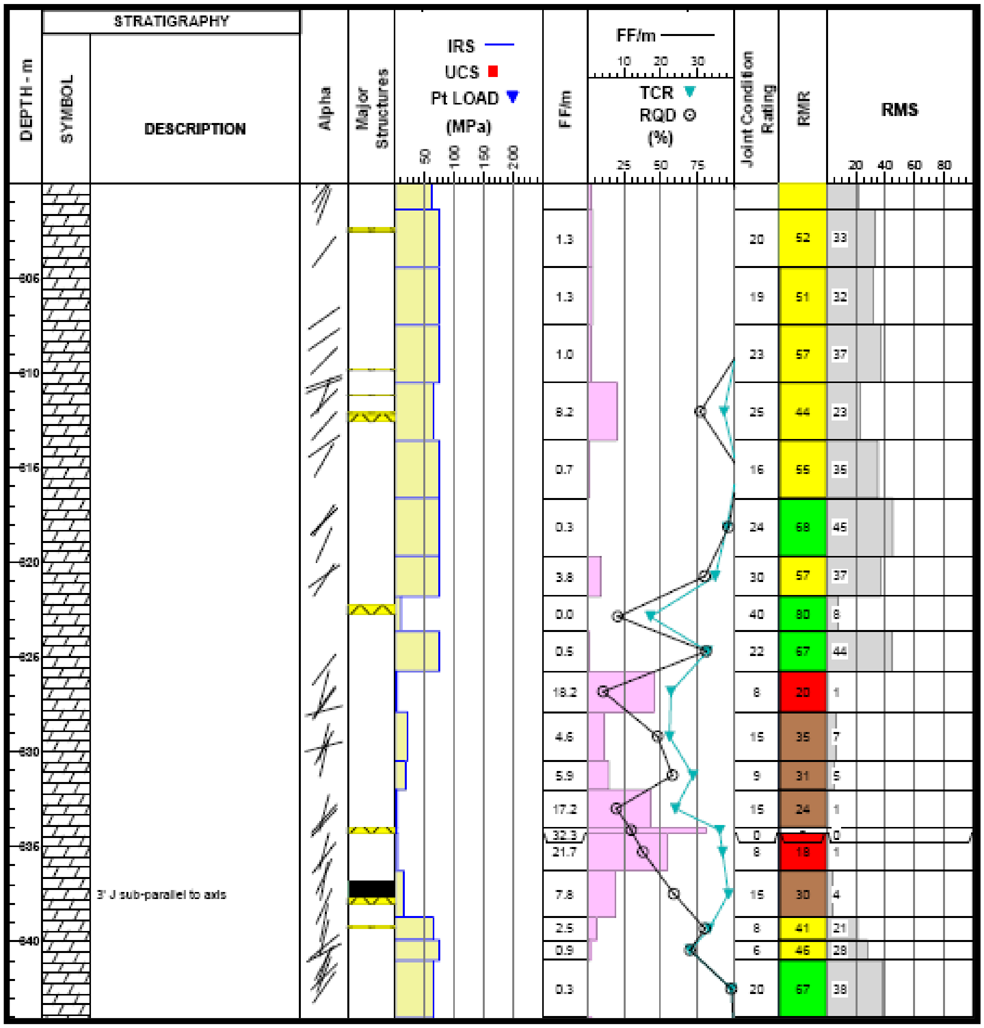


Figure 6: An example of a Geotechnical downhole QA/QC graph as produced by SRK Consulting using software developed by Sobek-Technologies.

### 5.3 *Conclusion*

There is not one solution that will solve all data collection situations but managing the strengths of different products can give an organization the competitive edge.

Many companies have a history of developing software systems in-house which, while fulfilling certain functionality, can become complicated as requirements evolve and develop a reliance on key personnel. Should these individuals leave the company, maintenance of the system becomes a concern. This can be avoided with a standard product from a well established vendor, like the acQuire data management system where the software is

## 6 REFERENCES


Chung, L. 2004. Microsoft Access or Microsoft SQL Server: What's Right in Your Organization? SQL Server technical article.

Rengers, N, Hack, R, Huisman, M, Slob, S & Zigterman, W. 2002. Information Technology applied to engineering geology. *In Engineering Geology for Developing Countries* – Proceedings of 9th Congress of the International Association for Engineering Geology and the Environment. Durban, South Africa – J.L.van Rooy and C.A. Jeremy, editors.