



By Dan Martin

Dan Martin is a Physical Scientist with the National Geodetic Survey, and he is the NGS Geodetic Advisor for the State of Vermont.

Databases

Are you a human database? Do you have one in your office? You know...the person who can recount most projects he has worked on over the last one, five, or even twenty years? He can remember the name of the marks that were set and the project specifics. When asked about a particular mark, he closes his eyes, lifts his head slightly, and slips into a narrative describing the mark and its surroundings in photographic detail. These are truly unique individuals that are an asset to any operation. But what happens when that person leaves? More often than not, his knowledge leaves with him and is never regained.

There is a natural tendency to rely on people like that who have a natural ability to store and retrieve information. Add to the mix that in today's world of technology, massive amounts of information can be generated in a relatively short period of time. Geodetic control that once took weeks or months to generate can be established with relative ease in a matter of days. The problem today is not generating the information, but keeping up with and managing the almost constant flow of information in and out of a survey operation.

So how do we protect ourselves when one of those "human databases" leaves? Short of plugging him into the office PC and downloading him, there is little we can do at the time of departure. However, with careful planning and determination to follow the plan, you can capture the collective memory of everyone in your office, and develop a truly powerful tool.

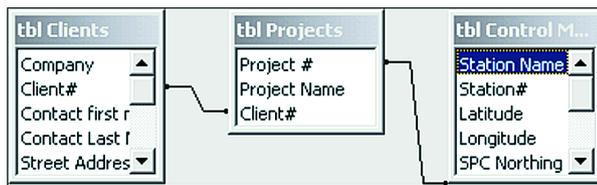
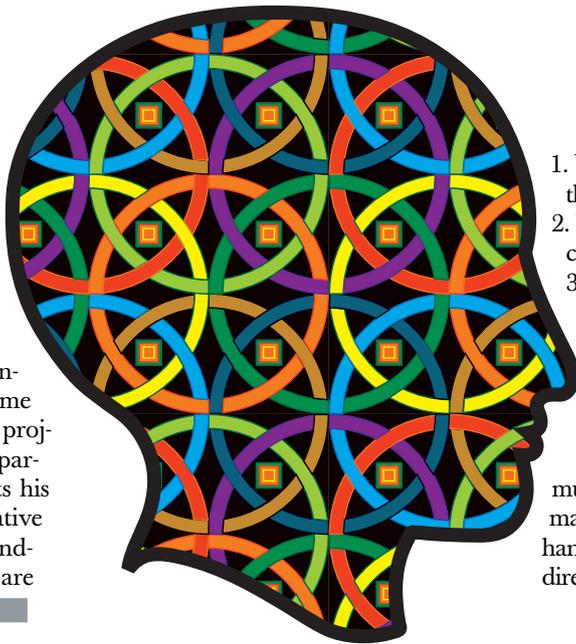


Figure 1 Linked relationships

Databases

A database, in essence, is nothing more than a standardized collection of information that is organized by categories. The power of a database can be realized through the ability to create queries and receive the desired answers. This of course can only be accomplished if the database is programmed with the information that is desired. Therefore, the development of a database requires significant thought and planning. Many people decide that developing and maintaining a database takes too much time. However, the initial time spent will be recovered many times over down the road.

Planning

When designing a database, ask yourself the following series of questions:

1. What types of queries will we submit to the database?
2. What information must the database contain to answer those queries?
3. Do we currently collect the information indicated in Question 2?
 - a. Yes. If so, what format is it in?
 - b. No. If so, can we collect it and what format should it be in?

Your desire should be to reuse as much data as possible in its native format and to minimize the amount of hand entry required. If data can be directly imported (in its native format)

into the database, great. If not, then look to see if the data can be reformatted into a compatible format for your database, e.g., a comma delimited text file.

Once you have established what data will reside in the database, you will need to

establish procedures for collecting and updating the data. This will require the establishment of standards or conventions that will need to be followed. All databases require standardization.

Garbage In/Garbage Out

In order to be useful, the data that resides in your database needs to be standardized. For example, let's say that I designed a database that contains control mark information, and one of the fields in that database is *Town Name*. I included this field so that I could extract all control marks in any given town. Now I go to my Super-Duper Database, and ask it a *simple* question. "Give me a list of all the control marks in the town of Saint Johnsbury." The database sorts for a second then spits out a listing of three control marks. I think to myself, "I remember setting six in that

town just last month!” Can anyone guess what the problem is? You probably guessed it; multiple spellings and abbreviations were used in the database for that town. On inspection of the database, we find the following entries in the *Town Name* field: Saint Johnsbury, St Johnsbury, St. Johnsbury, as well as the same permutations using Johnsbarre instead of Johnsbury. As you can see, the content of the database *must* be standardized in order to return the desired results.

Database Structure

Databases can be as simple as a single table of information, or more complex, containing several tables that are linked together. The type of data that you are storing and retrieving will determine the structure of your database. For example, a client database would most likely consist of one table named “Client Addresses”, whereas a project database might consist of multiple tables named *Projects*, *Client Addresses*, *Control Marks*, and *Project Maps*. Each of these tables would contain data fields relative to the table. For instance, the *Control Marks* table might contain fields named *Station Name*, *Station#*, *Latitude*, *Longitude*, *Horizontal Datum*, *Height*, *Vertical Datum*, *Project#*, and so on. The idea is that all information is only entered once, and can be accessed by linking like fields between tables. **Figure 1** shows how this linking relationship might look. This sample database contains three separate tables. Each table is designed to contain a certain category of information, but has been linked with other tables through the use of common fields. In this example, the *Projects* table and the *Control Marks* table both contain the *Project#* field. Since this is the case, a query could be designed to extract all control marks that resided in a particular project, and display their names and coordinates. And since the *Projects* table is linked to the *Client* table via the *Client#* field, the client information could be extracted at the same time. This example is fairly simple and may not have much practical use; however, it does demonstrate some of the simple functions of a database. A more practical use might be to search for existing control in a proposed project location. This could be done by setting criteria for minimum and maximum values in the northing and easting fields. Doing this would effectively establish a search box for the query.

What to Include

The type of information that you include will be related to the type of work that you do. Start by including just the information that you know you need, but be careful not to limit yourself by your definition of what data is. Data can be anything at all. For instance, data can be a file name. Think for a minute about the files you work with on a regular basis. Do you work with CAD files? How about digital photographs? The power of databases today allow you to store that filename in the database, as well as open the file from the database. If I have a table in my “Projects” database that contains a field named *Project Map*, then I could populate that field with the location and file name of my project maps, e.g., C:\projects\articles\ example.dgn. The database can be instructed that this field contains a file, which is opened with my CAD software. All I have to do is click on that file name, and the project map opens.

Seeing Is Believing

A much more effective way to use your survey database is to integrate it with GIS software. This enters the area of the “spatial database” and is truly a powerful data management and retrieval tool. The GIS software operates as a database, but has the additional capability to “show” you the results of your queries and the contents of your tables by using a graphical interface (as long as your tables contain a spatial component, i.e., coordinates). In addition to the data that you develop yourself, there are many products available that can be incorporated into a GIS database such as highway centerlines, town and county boundaries, railroad lines, parcel maps, and digital aerial photography. All of this information can be made accessible in the field through inexpensive (even free) GIS viewing software that can be loaded onto a crew’s laptop, allowing them direct access to the information in the field.

Making the decision (and commitment) to develop a survey database can be tough. But if you find yourself digging through more than two file cabinets or spending more than 10 minutes to compile a small amount of information or performing this type of task often, then it might be time to consider developing a database. *A*

Subscribe Today@
www.TheAmericanSurveyor.com