

BGS Geoscience Integrated Database System

A repository for corporate data

Garry Baker & Jeremy Giles, *Keyworth*

The BGS Geoscience Integrated Database System (BGS-geoIDS) project was conceived as a way to redress the widely recognised fragmentation and inaccessibility of corporate data and so prevent further loss of existing digital data. The main aim of the BGS-geoIDS project was to establish an integrated data

repository, which would contain digital geoscience data in its many forms. A secondary aim of the project was to provide the necessary tools to access the data in the most cost-effective way available to the BGS.

The project was planned to run over a three-year period, starting in 1998 and

completing in 2001. This length of time was necessary as the BGS has been collating data since its creation in 1835; digital data have been accumulating since databases and digital techniques were introduced into the survey in the early 1970s. This makes the BGS a truly data-rich organisation. To utilise fully the data stored and ensure that we maximise its potential, it is essential that we collate the digital data into a coherent form, the data repository. This single data store would provide BGS staff and the public with access to the entire BGS data holdings, subject to commercial-in-confidence clauses upon certain data bequeathed to the BGS from external organisations and companies.

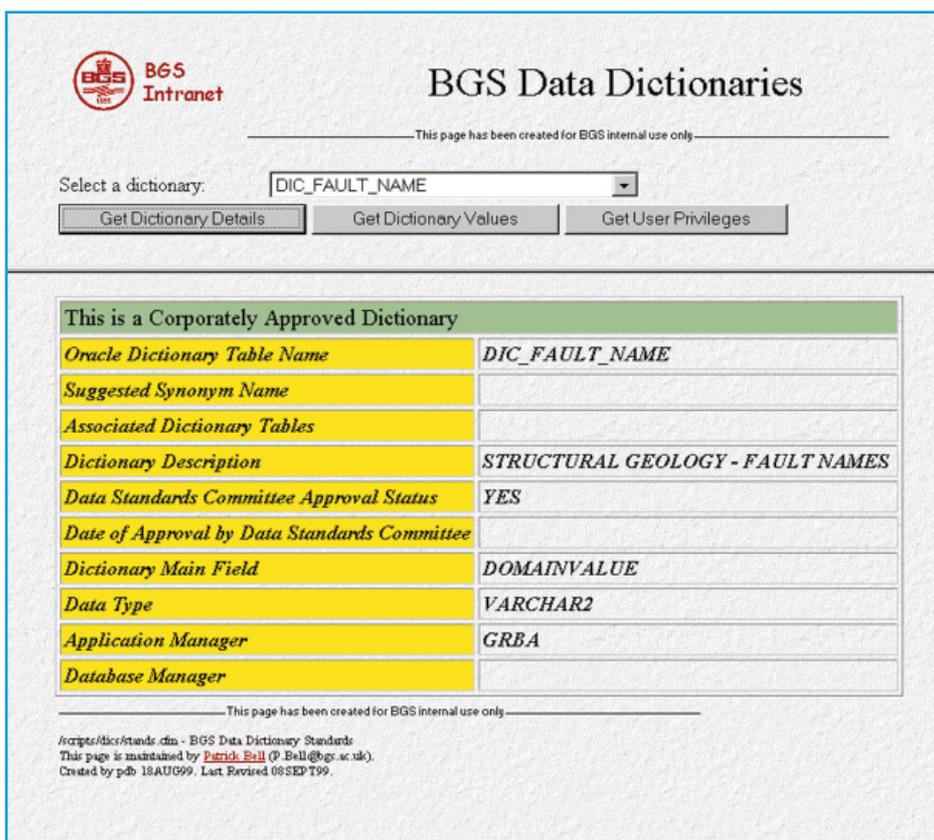
The project was focused upon nine main areas listed below.

- data model
- discovery metadata
- infrastructure management
- web-GIS Geoscience Data Index (GDI)
- Intranet Data Access (IDA)
- data standards (point and spatial)
- application standards
- external standards
- common data-sets

The nine main sub-projects can easily be thought of in four broader high-level categories. These are data model, Discovery Metadata, web-enabled querying of data, and standards.

Data Model

The data model is a map of the data tables we have in our corporate data repository. This map is termed an entity-relationship diagram by data modellers and allows the design and linkages of all the tables to be symbolised. This data model will ensure that all future new design, modification and query building carried out by BGS IT staff will take into consideration all the relevant data tables. Design modifications or upgrades will also be performed knowing the possible side effects that might occur to surrounding or linked tables within the repository.



This page has been created for BGS internal use only

Select a dictionary:

Get Dictionary Details Get Dictionary Values Get User Privileges

This is a Corporately Approved Dictionary	
Oracle Dictionary Table Name	DIC_FAULT_NAME
Suggested Synonym Name	
Associated Dictionary Tables	
Dictionary Description	STRUCTURAL GEOLOGY - FAULT NAMES
Data Standards Committee Approval Status	YES
Date of Approval by Data Standards Committee	
Dictionary Main Field	DOMAINVALUE
Data Type	VARCHAR2
Application Manager	GRBA
Database Manager	

This page has been created for BGS internal use only

/scripts/dicr/standr.din - BGS Data Dictionary Standards
This page is maintained by [Pamidi Pulli](#) (P.Pulli@bgs.ac.uk).
Created by gpb 18AUG99. Last Revised 09SEP199.

Discovery Metadata

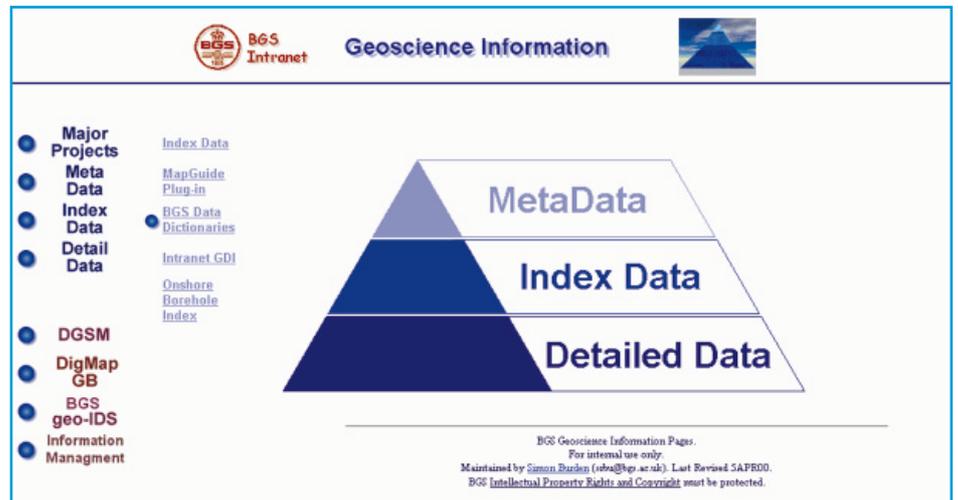
A data model of the data tables held is a useful tool for building queries to extract data from BGS tables but it does not intuitively inform the end-user what data we hold in the first place. This is the task of the BGS Discovery Metadata. Metadata is 'data about data' in its simplest form and allows us to query for 'data banks' (i.e. logically grouped data tables based upon a geoscience specialism or product). This metadata holds a description of the data bank, its usage, contact information, details on update frequency and logical keywords that relate to the data bank. An end-user searches the BGS metadata through a web-enabled interface either by typing in their choice of geoscience keywords or a direct data bank word search on the 'title' and/or 'description'.

“... to utilise fully the data stored it is essential that we collate the digital data into a coherent form ...”

Web-enabled querying of corporate data

A key component of BGS-geoIDS was recognised very early on in the project planning stages. A single data repository, mapped and searchable with structured metadata is a sound foundation, but will not succeed in significantly enhancing the BGS data holdings unless all our end-users can easily access the data they require. In this e-commerce age this means that web-enabled ways of querying data are required by customers and BGS staff. Web-enabled solutions have the benefit of being platform independent and easily deployable.

Two sub-projects within BGS-geoIDS have tackled these issues. The first is the Geoscience Data Index (GDI), both upon our Intranet and Internet sites. GDI is a spatial search engine that allows customers and staff to view the BGS's data holdings. The second sub-project is Intranet Data Access (IDA). This sub-project has built a web-enabled interface



Intranet Geoscience Information web page.

to the commonly accessed corporate data banks within the BGS, with the ability to query and manage them.

Standards

Standards are an essential component of all geoscience work within the BGS. We encounter them in applications, transfer formats, images, data dictionaries, and spatial dictionaries to mention just a few in the IT sphere. These have significant repercussions on the data repository and on queries carried out upon the data.

For example, to ensure that grid location information from a variety of differing data tables can be compiled we have to know the datum-line and type of grid used to compile each table. This requires us to adopt a standard locational grid and datum-line for our geoscience data, or to store within our repository selected and approved algorithms that will transform co-ordinates between differing grids and datum-lines.

This clearly demonstrates an example of the complexity of developing, managing and querying a large geoscience data repository. BGS-geoIDS has tackled and identified standards for applications, data dictionaries, dictionary design, spatial attributes and spatial dictionaries. These have been documented and can be found upon the BGS Intranet.

A project of this size and complexity has not been without problems and many lessons have been learnt along the way, but the completion of the project is rapidly approaching.

BGS-geoIDS have established a single Oracle® data repository, which is fully data modelled, documented and working towards a set of approved standards. A metadata engine that describes concisely the databanks we hold is available on both our internal and external websites, along with selected World Wide Web metadata gateways (such as the National Geospatial Data Framework, or NGDF). Finally, web-enabled applications have been implemented that facilitate easy search and retrieval of data held.

“... a single data repository is a sound foundation, but will not succeed unless all our end-users can easily access the data ...”

The BGS is in a much stronger position now than three years ago and is heading into the Digital Geoscience Spatial Model project (DGSM) with its data in good order.

Further details about the BGS-geoIDS project can be obtained by contacting:

Mr J R A Giles, Information Management Programme Manager, British Geological Survey, Keyworth, Nottingham NG12 5GG or E-mail: jrag@bgs.ac.uk