

BRITISH GEOLOGICAL SURVEY

INFORMATION Management PROGRAMME

INTERNAL REPORT OR/12/075

Information Management Focus and Operations: Delivery Plan 2011 - 2012 and Beyond

L Ault, P H O Henni, S F Hobbs, R R Lockett, T McCormick, A A
McKenzie and H M Glaves

Contributor

C E Morley

The National Grid and other
Ordnance Survey data © Crown
Copyright and database rights
2012. Ordnance Survey Licence
No. 100021290.

Keywords

Information Management, Data
Standards, Data Curation.

Bibliographical reference

L AULT, P H O HENNI, S F
HOBBS, R R LUCKETT, T
MCCORMICK, A A MCKENZIE
AND H M GLAVES. 2012.
Information Management Focus
and Operations: Delivery Plan
2011 - 2012 and Beyond. *British
Geological Survey Internal
Report*, OR/12/075. 40pp.

Copyright in materials derived
from the British Geological
Survey's work is owned by the
Natural Environment Research
Council (NERC) and/or the
authority that commissioned the
work. You may not copy or adapt
this publication without first
obtaining permission. Contact the
BGS Intellectual Property Rights
Section, British Geological
Survey, Keyworth,
e-mail ipr@bgs.ac.uk. You may
quote extracts of a reasonable
length without prior permission,
provided a full acknowledgement
is given of the source of the
extract.

Maps and diagrams in this book
use topography based on
Ordnance Survey mapping.

BRITISH GEOLOGICAL SURVEY

The full range of our publications is available from BGS shops at Nottingham, Edinburgh, London and Cardiff (Welsh publications only) see contact details below or shop online at www.geologyshop.com

The London Information Office also maintains a reference collection of BGS publications, including maps, for consultation.

We publish an annual catalogue of our maps and other publications; this catalogue is available online or from any of the BGS shops.

The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as basic research projects. It also undertakes programmes of technical aid in geology in developing countries.

The British Geological Survey is a component body of the Natural Environment Research Council.

British Geological Survey offices

BGS Central Enquiries Desk

Tel 0115 936 3143 Fax 0115 936 3276
email enquiries@bgs.ac.uk

Environmental Science Centre, Keyworth, Nottingham NG12 5GG

Tel 0115 936 3241 Fax 0115 936 3488
email sales@bgs.ac.uk

Murchison House, West Mains Road, Edinburgh EH9 3LA

Tel 0131 667 1000 Fax 0131 668 2683
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road, London SW7 5BD

Tel 020 7589 4090 Fax 020 7584 8270
Tel 020 7942 5344/45 email bgs_london@bgs.ac.uk

Columbus House, Greenmeadow Springs, Tongwynlais, Cardiff CF15 7NE

Tel 029 2052 1962 Fax 029 2052 1963

Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB

Tel 01491 838800 Fax 01491 692345

Geological Survey of Northern Ireland, Colby House, Stranmillis Court, Belfast BT9 5BF

Tel 028 9038 8462 Fax 028 9038 8461

www.bgs.ac.uk/gsni/

Parent Body

Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU

Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

Website www.bgs.ac.uk

Shop online at www.geologyshop.com

Foreword

This report provides an overview of the tasks undertaken by the Information Management Focus and Operations Team of the BGS for the year 2011 – 2012 and includes a forward look to future years.

Acknowledgements

Many individuals in BGS have contributed to work described here, several, from outside the Team are mentioned in the text. Acknowledgement is made to the contributions and hard work of all the staff who worked in the Information Management Focus and Operations Team during 2011 – 2012, to all those who contributed via strategic deliverables, with a specific thank you to Rob Armstrong, Tom Bide and Chris Royles, and to all data managers who realised the value of their data and took pains to organise and secure it for future users.

Contents

Foreword	i
Acknowledgements	i
Contents	i
Summary	ii
1 Introduction	1
2 Overview of Activities	1
3 BGS Wide Activities	1
3.1 Metadata, Data Management and Audit.....	1
4 Geoscience Themed Activities	7
4.1 Information Management Sector Managers – Senior Data Managers	7
4.2 Information Management in Geophysics 2011-12 and beyond (IM1).....	9
4.3 Information Management in Geology 2011-12 and beyond (IM2).....	10
4.4 Information Management in Hydrogeology 2011-12 and beyond (IM3).....	15
4.5 Information Management in Geochemistry 2011-12 and beyond (IM4).....	16
4.6 Coastal and Marine Information Management (IM6) 2011-12 and beyond	18
5 European Co-Funded Activities	22
5.1 Geo-Seas.....	22
6 Additional Tasks	26
6.1 Strategic Deliveables	26
6.2 Marine Data Management (Paul Henni).....	26
6.3 3D Fossil Scanning (Tim McCormick)	27

6.4	Marine Minerals (Tom Bide).....	27
6.5	Earth Observation Framework (EOF) (Rob Armstrong).....	28
6.6	Other activities.....	29
7	Conclusion.....	29
Appendix 1	INSPIRE.....	1
Appendix 2	Data Holdings Maintained by IMFO.....	2
	Glossary.....	4
	References.....	5

FIGURES

Figure 1	Data from the National Seismic Database being used to demonstrate to the media how an earthquake in Japan is recorded in the UK.....	9
Figure 2	3D Scan of an Ammonite from the Movie.....	11
Figure 3	Publically available BGS Sample Data locations.....	19

TABLES

Table 1	Disks making up the Keyworth SAN.....	2
Table 2	Information Management Sector Managers.....	7
Table 3	Sector Managers' Team Responsibilities.....	8
Table 4	Geo-Seas Outputs 2011 - 2012.....	22

Summary

This report provides an overview of the work undertaken by the Information Management Focus and Operations Team for the financial year 2011-12 and includes a forward look for future years.

Currently the Team's projects cover specialist data management activities under five geoscience themes, geophysics, geology, hydrogeology, geochemistry and marine plus a multi tasked project covering corporate metadata, proactive data verification and project data management. The Geo-Seas project, a Pan-European e-infrastructure project for the management of marine geological and geophysical data, is also managed under this Team.

1 Introduction

The Information Management Focus and Operations (IMFO) Team have the broad remit of promoting good data management within BGS by ensuring that project data are properly managed and preserved and can be used and potentially re-used by future researchers. A vital aspect of our work is to improve the quality of data in corporate databases, proactively by development of standard input procedures and reactively by systematic error correction within existing datasets.

The Team has responsibility for BGS Corporate metadata records, assuring project information assets, managing Storage Area Network (SAN) resources and overseeing the production and authorisation of Data Management Plans (DMPs).

The Information Management (IM) Sector Managers in each primary geoscience discipline have responsibility for the management and maintenance of corporate scientific databases in their sector. Liaising with the science Team Leaders they gain an understanding of developments and data resource requirements throughout the organisation.

In addition IMFO oversees and provides the co-funding for Geo-Seas a Pan-European e-infrastructure project for the management of marine geological and geophysical data (FP7-Infrastructures-2008-1.2.2).

2 Overview of Activities

Projects and tasks undertaken by the IMFO Team are:

1. Metadata, Data Management and Audit, including:
 - a. Management of SAN (Storage Area Network) resources.
 - b. Corporate Metadata
 - c. Proactive Data Verification
2. Geophysical Data Management
3. Geological Data Management
4. Hydrogeological Data Management
5. Geochemical Data Management
6. Marine and Coastal Data Management
7. Geo-Seas
8. Additional Tasks

3 BGS Wide Activities

3.1 METADATA, DATA MANAGEMENT AND AUDIT

Linda Ault

3.1.1 Management of the Storage Area Network (SAN)

A SAN is a high-speed network dedicated to information management and it consists of various levels of storage hierarchy that permit large volumes of data to be stored in a secure manner. For example, the Keyworth SAN storage (since extra storage capacity was added this year) is provided by 336 disks, divided into groups across 2 SANs:

Table 1 Disks making up the Keyworth SAN

HP SAN	
Disk Group	Number of Disks
1	40 x 300Gb FATA
2	48 x 300Gb FATA
3	80 x 300Gb FibreChannel
4	72 x 300Gb FibreChannel
Dell SAN	
Disk Group	Number of Disks
	96 x 2Tb SATA

Disk groups are divided into Virtual Disks (aka LUNs) which are presented to servers to make the space available to users/applications. The maximum size of a Virtual Disk is now 15TB whereas before the update the maximum size was 2TB.

The default size for a project folder is set at 250 MB; however projects can be allotted more space if there is a requirement for it. Several projects will reside on an individual Virtual Disk but for reasons of efficiency Virtual Disks should not be completely filled.

As for all disks, should a Virtual Disk become full no further data can be written to it. It is therefore important that SAN managers know how much space a project is likely to use so that the project folder is located where there is enough space for it. If projects on a Virtual Disk exceed their quotas to the extent that the disk becomes full none of the projects on that disk can save any further data. Given sufficient warning of the requirement for additional space projects can be moved between Virtual Disks, avoiding this situation. Problems arise when project staff are not aware of how much space they are using or if they ignore requests to advise SAN management staff of how much space they require.

A file with the name _README has been placed in each of the dataset folders of the corporate store (s: drive). The file will contain details of the Terms and Conditions for the dataset or a link to the person responsible for the data. Details of Terms and Conditions are necessary for legal and copyright reasons; failure to comply with them could lead to serious repercussions at both an organisational and individual level. It is vital that before using any data from the s: drive staff are aware of the terms and conditions of its use; whether it can be used freely, under licence or only for non commercial work.

Minor restructuring of the workspace area (w: drive) has taken place to accommodate changes in the matrix, new folders have been added for Science Facilities and some of the discontinued Team folders have been removed. There are more than 600 active project areas on the w: drive together with Heads of Science and Admin areas. The Data Management Planning System (DMPS) is used to manage the SAN space and to record a description of the data held in each folder, the terms and conditions and datasets produced; recording the information necessary for the discovery of research assets at project closure.

3.1.2 The Digital Archive

The Digital Archive, a tape storage system for data less frequently accessed, was updated in collaboration with Records Management and Database Solutions staff. The look is now similar

to the Data Management Planning System. There are three main sections: Create, Search, and Reports.

The Create section enables the user to create an archive request. The request will involve the creation of archive metadata to describe the data being archived.

The Search section enables users to search for particular archives or files within archives. A number of searches are available; archive number, text string, archive requester, file path etc.

The Reports section provides a number of searches as reports listing such things as, current archives, archive requests pending authorisation, archives requested by the user etc.

There are two new pieces of functionality, the Vault and Review.

The Vault is a special area of the SAN disk space set aside for read only data. It is visible in Windows as the V: drive. Files from completed work that require quick access for reference, operational or legal reasons can be stored in the Vault in the short term. Files in the Vault will be subject to review on a defined timescale, currently one year, with the aim of either disposal or formal archiving to tape.

Review capability enables Records Management staff to review archive records for continued relevance or for continued confidentiality. No changes will be made by Records Management staff to archived data without consultation with the data owner.

3.1.3 Data Transfer Area

A Data Transfer Area W:\DataTransfer has been set up to provide for the secure storage of data received on disk, tape etc. when project space allocations are inadequate. This folder can be used for temporary storage of datasets prior to storage in the Digital Archive. A procedure for the scheduling of the use of this space is available via the Digital Archive intranet pages. This storage area is for temporary use only, it is not backed up. Supply media should not be wiped or returned until secure storage in the Digital Archive has been verified.

3.1.4 Future Tasks – Data Management

Changes to RMS/Project codes (one code per Team, one task per project) has meant a rethink on how project folders can be assigned. It is vital that SAN folders relate to specific items of work with data deliverables that can be garnered for long-term use and re-use.

Datasets produced by projects need to be assessed for significance and stored appropriately when the project comes to an end. This may mean transfer to the digital archive, storage on the s: drive, in corporate databases or (for hard copy) the National Geoscience Data Centre. Systems to make this a simple and routine process need to be established; i.e. “taking away the disincentives” to data handover and long-term data centre storage.

3.1.5 Metadata Profiles and Data Discovery

The project is responsible for collecting metadata on significant datasets, entering it into the corporate metadata tables and ensuring relevant records are available on the BGS website, and other services.

The metadata used to describe datasets; the “Discovery Metadata” conforms to the standard BS ISO 19115:2003, however it can be extracted from the database and displayed in various formats as required. BGS has to supply information on its datasets (or subsets of them) to (among others) the NERC Data Catalogue Service and data.gov.uk. Members of the IMFO Team work with staff from Database Solutions and Web Systems to ensure metadata compliant with relevant standards (INSPIRE, UK GEMINI) is supplied to the necessary distribution systems for further and wider onward dissemination.

This work has entailed making sure that information displayed on the services is a true representation of the information held; some values having to be aggregated or consolidated to comply with required formats.

3.1.6 NERC Data Policy - Data or Product?

The NERC Data Policy has been revised; the updated policy came into force in January 2011 with specific requirements for data management plans held back until 2012. The policy states:

“All the environmental data held by the NERC Environmental Data Centres will normally be made openly available to any person or any organisation who requests them. “

It became necessary to clarify which datasets within the organisation were classified as data and therefore freely available and which could be classified as products to be sold or licensed. The IMFO Team together with members of Information Products made an initial assessment of all registered datasets which provided the basis for the formal classifications to be made (Wildman, Westhead, Ault, Henni and Campbell 2011).

3.1.7 Technical Metadata

Technical metadata describes the objects that store the data within a database management system and the objects' relationships within the system. The specific Technical Metadata discussed here are those used to describe the objects in the BGS Oracle system. These Metadata are held in Oracle tables which can be searched via the Technical Metadata Application on the intranet accessed via the Geoscience home page.

Most of the work on technical metadata updates during the year was concerned with obtaining descriptions for databases, tables, views and triggers amongst others. Obtaining descriptions relies on knowing the identity of the table manager; in some cases merely by inference. There are many objects where no data manager has been recorded and the task of allocating managers (and to a lesser degree application managers) is vital. In many cases this is not easy. Visiting people in person produces a much more positive and successful response than contact by email. This work is done under the guidance of the BGS Data Architect with the Database Solutions Team.

3.1.8 Future Tasks 2012-13 – Metadata

Discovery metadata for academic datasets received by the National Geosciences Data Centre is entered into the Earth Science Academic Archive collection in the metadata tables. Currently this metadata collection is only visible on the BGS website and is not available (at record level) via other portals such as the NERC Data Catalogue Service. Work has begun to verify data in this collection and prepare it for wider distribution.

Project data assets often include datasets which should be entered into data centre holdings when the project ends (as per the updated NERC Data Policy). Metadata must be provided and displayed for these datasets so that they can be “discovered” by any potential users both within and outside of BGS. Workflows are being established to facilitate this. Aspects of the workflow will rely upon amendments being made to metadata and data management systems.

The discovery metadata system for GIS metadata, developed by Andrew Marchant, Graham Smith and Rachel Heaven consists of a customised data entry system accessed via ArcCatalog when the GIS datasets are being created. This system populates the corporate ISO compliant metadata tables with the information entered. This metadata collection needs to be further integrated within the discovery metadata system and applications so that it can be harvested and displayed as for other BGS datasets.

Plans for technical metadata include extensions to the system to fully integrate it with the Discovery Metadata System at database level. The architecture is in place and population and enriching of the content of the corporate databases metadata table is going ahead. The Data Architect expects that this year (2012-13) the link between Technical and Discovery will be

more formal and BGS will have a definitive list of corporate databases associated with corporate datasets.

Other plans are to make more of the technical metadata visible including, potentially, publishing the objects externally.

3.1.9 Data Verification

Data verification has an increasing role within the organisation. Government open data initiatives mean there is a greater emphasis on guarding against potential reputational and litigational damage. In addition to this, securing significant project data for use and re-use by future researchers dictates the following of standards to promote interoperability and the building of networks.

Verification work aims to improve the quality of data in corporate databases, proactively by development of standard input procedures and reactively by systematic error correction within existing datasets.

3.1.10 Verification - Geoscience Issues Log

The issues log is the primary corporate tool for reporting perceived errors in corporate data holdings. Issues with data are categorised and assigned to appropriate staff (of a small dedicated team) for resolution. The success of many of BGS's products and services rely on corporate data and hence, at the most fundamental level, does the image of BGS itself. Ensuring that issues reported are dealt with speedily and effectively is core to the perception of BGS as an organisation setting great store on the quality of the data that underpins its activities and products.

The release of borehole scans on the BGS website has meant that external customers can now report any perceived issues with borehole data and these are dealt with as routine. Borehole scans feature in the most visited sites on the BGS website. Three related databases are involved in a workflow from data receipt to release for viewing. Verification tasks demand an understanding of all the processes that contribute towards ensuring the scans are indexed and viewable.

New work this year has included working with Wallingford staff on updates to WellMaster records; typically records that are associated with invalid SOBI entries.

The scale of the reduction in errors is impressive, 72% of errors in Borehole Geology (some of long-standing) and 50,000 updates to SOBI. These two datasets are the foundation for much of the current modelling work in BGS. Understanding, conceptualisation and modelling of ground conditions can only occur with reliable borehole information.

3.1.11 Verification – Single Onshore Borehole Index (SOBI)

SOBI (Single Onshore Borehole Index) is one of the databases in the complex workflow that leads to borehole scans online. It has a complex history, evolving over time from its precursors. Understanding database business rules and historical decisions is vital if systems are to be stable and sustainable, paving the way for future development and enhancement.

A first draft of the SOBI handbook was compiled in collaboration with members of the Records Management Team. This report documents the history of SOBI and outlines current procedures for the registration of borehole records. It is a reference document of database content and procedures; useful to the registration team but also for users interested in gaining an understanding of how SOBI works (Bowie, Gillanders and Morley, 2012).

A review of the dictionaries used in SOBI was undertaken to determine which codes are currently in use and also to determine whether these existing codes are adequate to cover all current requirements.

The SOBI verification team, meet on a monthly basis to review SOBI issues and documentation.

3.1.12 Verification - Borehole Geology

There has been a concerted effort to work systematically through the depth difference errors on the Borehole Geology database; 2500 errors corrected up to press.

21,000 of the errors automatically flagged by error checking routines have been corrected, some by global updates but a considerable number by individual intervention by members of the verification team.

3.1.13 Verification - Other Activities

Work has started on the systematic checking and correction of errors involving site investigation locations and their associated borehole records.

Advice on Borehole Geology business rules was provided to aid the development of a new Borehole Geology data entry system (Porcupine).

3.1.14 Future Tasks 2012-13 – Verification

Work will continue on correcting errors identified in the Geoscience Issues Log; typically for SOBI and Borehole Geology. This is particularly important because of current high profile work on 3D data modelling products and services.

The Borehole Tsar role (providing an overview of borehole related activities and point of contact for borehole information) has been expanded to include expertise from Records Management, geological data management and data verification.

It is planned to add application issues to the Geoscience Issues Log; setting standard procedures where possible and tracking issue resolution.

The verification team will work closely with the Information Sector Manager for geology to begin systematic verification work on the Lexicon of Named Rock Units.

3.1.15 Future Tasks – Communication and Training

3.1.15.1 COMMUNICATION

Staff often experience difficulty locating applications for processes that are used infrequently. Data management tools are available on the intranet but in several locations based on their use. IMFO is producing a central Information Management page linking into all IM assets; this way staff will only have to rely on remembering the location of one page. This page will also be linked into the BGS Management Toolkit on the BGS intranet. This toolkit covers the myriad management activities that staff should be aware of, from Health and Safety to Roles and Responsibilities, Training, Project Management and Merit Promotion.

3.1.15.2 TRAINING

There are two main requirements here:

1. Data management policies and project responsibilities.
2. Use of existing applications.

BGS staff need to be aware of their responsibilities under the Freedom of Information Act and Environmental Information Regulations. In addition, open data is becoming increasingly more important. BGS is mandated by Government, via the Research Councils to make research outputs accessible for future use by other researchers. This means knowing how to store data routinely so that the information preserved is independently understandable to the whole Designated Community. Low awareness of data curation and preservation issues can lead to data

loss and reduced productivity. Recovery of significant research data from records set aside for even a year takes considerably longer than working on the data as soon as it is produced. This is a particular challenge when the data creator has retired or moved to another institution.

Staff need to be aware of the systems and help available to them. The proposed training modules will cover demonstrations of existing data management applications, e.g. the Digital Archive, Intranet Data Access and GeoIndex. There will also be links to relevant policies and staff contacts who can provide assistance covering things as diverse as record retention schedules and how to enter borehole geology data into the corporate database.

4 Geoscience Themed Activities

4.1 INFORMATION MANAGEMENT SECTOR MANAGERS – SENIOR DATA MANAGERS

Currently data management activities are divided into five sectors, each sector being the responsibility of the senior data manager for that area of science.

Table 2 Information Management Sector Managers

Sector	Manager	Manager's Location
Geophysics	Richard Luckett	Edinburgh
Geology	Tim McCormick	Keyworth
Hydrogeology	Andy McKenzie	Wallingford
Geochemistry	Susan Hobbs	Keyworth
Marine Geology	Paul Henni	Edinburgh

Although nominally representing their area of science, the Sector Managers will also provide local data management support for scientists at their office location as required.

4.1.1 Main Activities of IM Sector Managers

The IM sector managers have the responsibility to manage BGS project research data, vital to BGS commitments and reputation. They ensure they are gathered, stored, maintained and enhanced for ongoing use and re-use. The work covers hydrogeology and seismology to geology, geochemistry and geomagnetism. Tasks include database maintenance and enhancement, population and verification. Selected dataset specific management procedures are developed and released as operational.

Databases maintained include, but are not limited to: Borehole Geology, BritPits, Coastal and Marine Database, Data Management Planning System, Discovery Metadata, Geochemistry Database, Geoscience Imagebase 2 and Geoscentic, Geoscience Issues Log, Lexicon of Named Rock Units, National and Global Geomagnetic database, National and Global Seismic database, National Building Stones database, National Geotechnical Properties database, National Groundwater Level Archive, National Well Record Archive/WellMaster, World and UK Mineral Statistics.

Sector Manager duties include:

1. Advising Project Leaders on Information Management.
2. Ensuring projects produce adequate Data Management Plans.
3. Being the second signatory on Data Management Plans. ("Checked by")
4. Ensuring dataset producers provide Discovery Metadata.

5. Advising on data retention and disposal at project closure (see Information Management procedure 1.1.2 Project Asset Assessment and Management).

One of the important tasks for the Sector Managers is to meet with BGS Team Leaders to discuss requirements for the types of data being produced/collected, finding what the issues are and what future storage requirements/systems are likely to be needed.

To this end, the BGS Teams have been assigned to a specific IM Sector Manager in the first instance. Some BGS Teams will have overlaps between more than one IM sector but there is a nominal first contact for each Team, this contact will be the Sector Manager who checks relevant Project DMPs and liaises with the Team Leader. This year, Team responsibilities have been mapped as follows:

Table 3 Sector Managers' Team Responsibilities

4.1.1.1 GEOPHYSICS

- Earthquake Seismology
- Geomagnetism
- Volcanology
- CO2 Storage
- Advanced Seismic Techniques
- Geophysical Tomography

4.1.1.2 GEOLOGY

- Palaeoclimate & Palaeoenvironment
- Renewables & Energy Security
- Geology & Landscape NI
- Energy & Minerals
- Geology and Landscape England
- Chief Geologist Scotland
- Chief Geologist Wales
- Geotechnical & Geophysical Properties & Processes
- Shallow Geohazards & Risks
- Radioactive Waste
- Mathematical Modelling
- Geological Modelling Systems
- National Geological Model
- Earth & Planetary Observation & Monitoring

4.1.1.3 HYDROGEOLOGY

- Environment & Hazards
- Groundwater Systems
- Groundwater Processes
- Groundwater Protection
- Urban Geoscience

4.1.1.4 GEOCHEMISTRY

- Soils, Landscape & Climate Processes
- Geochemical Baselines & Medical Geology
- Minerals Sustainability

Mineral Resources & Policy

4.1.1.5 MARINE GEOLOGY

Coastal & Fluvial Environmental Change

Marine Geology

Marine Geohazards & Geotechnology

Continental Margins

Regional Hydrocarbon Prospectivity

4.2 INFORMATION MANAGEMENT IN GEOPHYSICS 2011-12 AND BEYOND (IM1)*Richard Luckett***4.2.1 National and Global Seismic database**

A year of continuous waveform data for the stations in the national network has been added to the database. Four more stations were added to the network this year and integrated into the data acquisition and quality control systems. In addition several years of continuous data have been added to the database from a small array of broadband seismometers set up by Cambridge University as part of a PhD project. The World Seismic Database of earthquake parameters has been modernised this year with the unified magnitude being changed and automatic procedures introduced to keep it up to date with national and global events. Monthly and annual reports were produced on time.

Figure 1 Data from the National Seismic Database being used to demonstrate to the media how an earthquake in Japan is recorded in the UK

**4.2.2 National and Global Geomagnetic database**

Geomagnetic data from BGS observatories in the UK and overseas has been added to the data base for the year, including continuous time-series data and derived statistical products. Monthly and annual reports were produced on time.

4.2.3 Seismology and Geomagnetism web services

The separate Seismology and Geomagnetism web servers in Edinburgh have been updated and reformatted to conform to corporate website appearance. They continue to be updated and maintained to a high level of reliability and resilience.

4.2.4 RECALL

The RECALL system continues to be the corporate resource for storage of geophysical log data. IM1 supports the maintenance of RECALL and the loading and verification of new data.

4.2.5 Geophysical databases

IM1 continues to support the maintenance of legacy corporate databases which are assets of the National Geoscience Data Centre. These include the Stratigraphic Surfaces Database, the Seismic Locations and Sections database and the Seismic Line Index database.

4.2.6 Other Tasks in IM1

Coal Authority metadata: IM1 supports the generation of metadata and its input into RECALL for the remaining, poor condition, Coal Authority data on an 'as is required' basis.

4.2.7 Future Tasks 2012-13

Support from IM1 has been requested for the new Global Volcanic Model – an Oracle database of volcanic behaviour that will be contributed to by researchers around the world and hosted by the BGS.

4.3 INFORMATION MANAGEMENT IN GEOLOGY 2011-12 AND BEYOND (IM2)

Tim McCormick

4.3.1 Project Management and Miscellaneous Geological Data Management FY 2011-12

This task includes provision of advice on geological information management and archiving across BGS, through informal interviews and other contacts with Team and Project leaders; also enforcement of information management responsibilities on projects, through assessing and approving project data management plans and SAN quota requests.

'One-off' activities to improve data quality in corporate databases this Financial Year have included validating and uploading approximately 12,400 metadata records describing fossils in the BGS museum collection that were on spreadsheets but had never been entered onto *Palaeosaurus*, as well as various other data management operations in that database. It is worth noting that the success of *Palaeosaurus* was instrumental in enabling the Collections Manager to obtain funding from JISC to work with Oxford University Museum of Natural History, Sedgwick Museum of Earth Sciences Cambridge, National Museum of Wales and the Natural History Museum to create a web database of images and 3D scans of British type fossils.

In August IM2 made a 4 ½ minute long movie using 3D scans of selected fossils from the BGS collections which had previously been captured under a Strategic Deliverable project. This movie was displayed at the World Conference on Marine Biodiversity in Aberdeen, 26-30 September. It is now available as part of Discovering Geology on the BGS website. In addition, a request has been received from the Scottish Fisheries Museum, Anstruther, Fife, to use it in a public gallery display, and permission has been given.

Figure 2 3D Scan of an Ammonite from the Movie

The BGS Linked Data Workshop held on 11 November highlighted the importance of IM2; every BGS example and pilot study presented throughout the day relied on data content produced, maintained, and owned by IM2: The Lexicon of Named Rock Units, geochronology dictionaries, rock and sediment type dictionaries, numerous other dictionaries and vocabularies – referred to in one presentation as "mature and respected vocabularies".

4.3.2 Project Management and Miscellaneous Geological Data Management FY 2012-13

Provision of advice and operation of information management procedures will continue, along with identification of data sets to be stored corporately, moving them to the S-drive or the Oracle database as appropriate and creating metadata and data management plans.

The implementation of the new NERC data citation policy on selected data sets will be trialled.

4.3.3 SOBI & Borehole Geology

Borehole Geology is the database of geological interpretations of boreholes drilled in Great Britain, supplied by geotechnical companies or made by BGS geologists or other agents. It is heavily used in conjunction with the Single Onshore Borehole Index (SOBI) to generate 3D models and other representations of the geology of Great Britain. Borehole Geology passed the 3 million records milestone in June 2011 and at time of writing holds over 3.1 million records.

FY2011-12: The never-ending work of data validation and improvement, and promoting coding standards has continued. Seven 'content codes' for new borehole geology coding projects have been assigned. In October 36,054 new records from the National Geotechnical Properties Database (see below) were copied into Borehole Geology, making them available for use in 3D modelling and other applications.

FY 2012-13: IM2 will continue data validation and providing assistance to geologists using the system. The project will develop and trial a more efficient system for copying records from the National Geotechnical Properties Database into Borehole Geology, and increase automation of checks on Borehole Geology to further improve confidence in the data.

IM2 will design proper interaction between Borehole Geology and the new ways of accessing and editing it. These include GSI3D and the Porcupine borehole viewer/editor. It is important to ensure the database meets the evolving requirements of geoscientists, but also to ensure the safety and integrity of the data.

4.3.4 Lexicon of Named Rock Units

This is the database of definitions of rock units and other geological deposits recognised by BGS on the UK and its continental shelf. As well as being a repository of knowledge, the Lexicon is used to constrain the lithostratigraphical attribution of other BGS information systems, including DigMap, 3D models, GeoReports, Geology of Britain Viewer, iPhone app, and many more. It holds approximately 19,500 records. It is searchable on the BGS website where it is regularly in the top 10 most viewed pages. A redesigned database offering enhanced functionality will be released internally by 1 May 2012.

FY 2011-12: An article by Tim McCormick describing the Lexicon appeared in Proceedings of the Geoscience Information Society in June 2011.

Also in June, as a result of BGS Opportunities Fund money, digital versions of 7 volumes of 'Lithostratigraphic Nomenclature of the UK North Sea', originally published by BGS in conjunction with the UK Offshore Operators Association (UKOOA) between 1992 and 1994, were made available for download from the BGS website.

IM2 subsequently commissioned BGS Publishing to produce digital versions of 'Stratigraphic Nomenclature of the UK North West Margin' (3 volumes, originally published 1996-99) and 'Lithostratigraphic Nomenclature of the Triassic, Permian and Carboniferous of the UK and offshore East Irish Sea Basin' (1 volume, originally published 1996). These will be loaded to the BGS website and information will be extracted into the Lexicon. IPR clearance has been secured from the relevant contributor oil companies.

In January 2012 the stratigraphical framework for the Quaternary of the UK continental shelf was set up in the Lexicon, following publication of BGS Research Report RR/11/03.

FY 2012-13: IM2 will validate and improve content to ensure a consistent baseline of completeness for all units on the BGS UK Stratigraphy Charts, and continue to update the Lexicon with reference to the BGS published onshore and offshore lithostratigraphic framework reports. This is a major task that will continue into succeeding years. IM2 will also seek to establish communications with users and other stakeholders outside BGS to inform and target future development.

4.3.5 BritPits

Also known as the mines and quarries database, this contains over 156,000 entries for active and inactive onshore mineral workings, as well as rail depots and wharves involved in mineral commodities movement. These data are used by BGS Minerals and Waste to produce the 'Directory of Mines and Quarries'. They are also used for market intelligence and analysis, and resource planning by BGS and by customers including local councils. The data are made available under licence.

FY 2011-12: The latest edition of the Directory (9th Edition) was published in April 2011. Data verification, loading, update and validation have continued throughout the year.

FY 2012-13: As well as ongoing data management it is anticipated that there will be minor modifications made to some of the queries and reports within the database.

4.3.6 World and UK Mineral Statistics

This is a time of heightened interest in mineral supply security. This database is maintained and updated by IM2, while BGS Minerals and Waste use it for market intelligence and analysis, and to produce ‘World Mineral Production’, ‘European Mineral Statistics’, and ‘UK Minerals Yearbook’ which provide reliable information to policy makers, regulators, industry and the public.

FY 2011-12: UK Minerals Yearbook was slimmed down and released as web-only, to reduce costs because of the end of the 5-year BGS-Dept. for Communities and Local Government (DCLG) Joint Minerals Information Programme. IM2 staff had to learn a slightly different production mechanism, with the book being formatted by Minerals rather than BGS graphics. “United Kingdom Minerals Yearbook 2010” was published on the MineralsUK website in July 2011. Data management continued ready for web-publication of the next edition which took place at the beginning of May 2012.

‘World Mineral Production 2006-2010’ was published in February 2012, and ‘European Mineral Statistics 2006-2010’ was released on the website on 4th April. European trade figures arrived in a different format to previous years which caused some difficulty.

Minor modifications have been made to the structure and user interface of the database.

IM2 staff contributed data to the updated BGS Mineral Profile for Rare Earth Elements, published on the MineralsUK website in November 2011, and to the forthcoming Critical Metals Handbook to be published by Wiley-Blackwell in autumn 2012.

BGS Minerals and Waste had discussions with Granta Design in Cambridge who make software used by manufacturing companies to assess supply risks due to political, environmental and other causes. Granta are interested in using BGS mineral production data. BGS Minerals and Waste are directly responsible, but the opportunity relies on data input and maintained by IM2.

FY 2012-13: as well as continuing data management, IM2 will begin to load available historical commodities data to the database, to allow analyses in a historical context. This year it is hoped to make a significant start on this, by taking data from old editions of BGS report series including UK Minerals Yearbook (which goes back to 1986) and UK Mineral Statistics (1973-1986). The database will also be extended to hold long time-series data for the UK which is currently held separately, and to enable xls-format downloads for registered enquirers.

4.3.7 Geotechnical Properties Database

This is a database of geotechnical properties and other information derived from site investigation reports supplied to BGS. It is used in engineering geology research, answering enquiries, and to attribute 3D geological models. It currently holds approx. 90,000 borehole records plus related data.

FY 2011-12: Data input and validation has continued throughout the year. In October, 36,054 records from this database were copied into Borehole Geology, making them available for use in 3D modelling and other applications.

FY 2012-13: IM2 will develop and trial a more efficient system for batch-copying records from the National Geotechnical Properties Database into Borehole Geology. The project also hopes to develop and trial a more streamlined workflow for receiving and loading data in the digital AGS (Association of Geotechnical Specialists) format, in collaboration with the Database Solutions Team.

4.3.8 National Building Stones Database

The National Building Stones Database holds data on (mostly historic) buildings and construction materials. The information is collected and supplied by English Heritage and Historic Scotland staff. BGS uses the information to answer enquiries and provide advice to

customers and partners, often local authorities and heritage agencies. There are approximately 4,000 index records plus related data.

FY 2011-12: Fine tuning of the database queries and tables has been done, and data population has continued throughout the year, co-funded by BGS and Historic Scotland. A research fellow (previously at Historic Scotland) has now joined BGS, but will not receive funding from IM2.

FY 2012-13: Next financial year data verification, entry and validation will continue, and there will be some minor modifications made to the database to aid edit-ability.

4.3.9 Dictionary Management

BGS makes extensive use of ‘dictionary’ database tables to constrain its domain vocabularies and promote standardisation and interoperability between its databases. IM2 gathers requirements and populates, promotes and maintains the key geological dictionaries, including those for rock and sediment types and chronostratigraphy.

FY 2011-12: Much of this year’s work has been of the nature of ‘business as usual’ maintaining dictionaries, for example producing new rock-type codes to be used in DiGMap and map production as needed. Major developments included a new graphical chronostratigraphy viewer / picker, and an intuitive lithology dictionary search tool, both released on the IDA.

In December a long-standing problem with the representation of Unlithified Deposits Coding Scheme sediment types in Oracle, in which qualifiers had been populated incorrectly in some dictionary entries, was fixed. This needed to be handled carefully with appropriate liaison with key users, because the possibility existed that some users had been misled about the meaning of some codes in the past, so that correcting this problem could potentially change the meaning of some records in databases.

FY 2012-13: IM2 will continue to maintain, rationalise and develop dictionaries and other controlled vocabularies as required to support BGS geoscience information systems. The project will trial adding definitions to dictionaries, including situations where the definition varies depending on the context in which the term is being used. We will rationalise the rock- and sediment-type dictionaries, removing obsolete ones that are still in use, encouraging users to move to the new up to date versions.

4.3.10 Geoscience Imagebase 2 (GI 2)

This is the BGS database of digital geoscientific images and associated metadata, and includes pictures taken by BGS geoscientists as well as collections accessioned from other organisations. A subset of the images are selected and published on the BGS website as ‘Geoscentic’, which is consistently one of the most popular areas of the BGS website.

FY 2011-12: Over the current financial year two staff have been trained to take over from the experienced staff who moved to the BGS Library at the end of the last financial year.

Work has continued throughout the financial year with projects across BGS to prepare their images for upload and to catch-up on the backlog of photographs awaiting upload.

In October 2011 Caroline Adkin gave a presentation on Geoscience Imagebase and Geoscentic at Wallingford and kicked off the process of Wallingford providing their digital images and metadata for upload. David MacDonald has agreed to act as conduit for communication between the GI2 personnel and Wallingford, and will be given a small allocation on IM2 from next FY to facilitate this.

FY 2012-13: Next financial year IM2 hope to trial user self-load of images, with a full roll-out later in the year if the trial goes well. The project will also develop categorization of images, with input from domain experts and projects. This is already underway with respect to geohazards and volcanoes. IM2 will also continue to help projects upload their images, and accession external collections.

4.3.11 IUGS-CGI Multilingual Thesaurus of Geoscience (MTG)

The MTG working group was instigated under the auspices of the IUGS-CGI (Commission for Geoscience Information) in 2003 to revise and improve their existing multilingual vocabulary. The multilingual thesaurus will be made freely available and has the potential to become an essential tool for indexing and searching, and to enable semantic interoperability across language boundaries. IM2 represents BGS on the CGI MTG working group.

The CGI took the decision in late 2010 to merge 2 of their task groups: the MTG working group and the CGI Concept Definitions Task Group. The work programme for the new group is in the process of being developed, but the membership of the new merged group has not been finalized.

FY 2011-12: In October 2011 the MTG collaboration application, used by the MTG group members to construct the thesaurus, was moved onto the NERC Extranet for improved application security.

FY 2012-13: Rachel Heaven is going on maternity leave in February 2012. This may signal the end of IM2 representation on the working group, and it is unclear what representation BGS will have on the replacement merged group.

4.4 INFORMATION MANAGEMENT IN HYDROGEOLOGY 2011-12 AND BEYOND (IM3)

Andy McKenzie

4.4.1 Hydrogeological records and WellMaster

FY 2011-12:

Normal acquisition registering and scanning of water wells continued, with attention being paid to resolving issues of scan quality reported through the issues log. Some staff resource was diverted to ensuring that public water supplies were easily identifiable.

WellMaster data entry focused on newly acquired data, and infill of 10 km squares where data entry percentages are low (usually because they were subject to a quick pass early in digitisation).

Data conditioning work examining primary aquifers and geological details for consistency was completed. Further classification of aquifers is possible, but requires significant hydrogeological input, so is now proceeding only where required by research projects.

The next phase of work was commenced, examining yield and water level data.

FY 2012-13

Water well data acquisition will continue.

If a decision is taken to release WellMaster data on OpenGeoscience, the data conditioning will need to be prioritised, and consideration given to developing clear flags for potentially anomalous data, and suitable tools to allow user feedback of errors and omissions.

Subject to resourcing, work may be done to check that all boreholes recorded in published well and spring catalogues prior to 1930 are correctly registered and available as scans.

4.4.2 Water level data

FY 2011-12:

Groundwater level data continued to be acquired from measuring agencies, and used to produce monthly hydrological summaries in conjunction with CEH.

Telemetered water level data is now received automatically from three research sites.

New pages featuring live graphical time series and enhanced descriptive material for key boreholes have been prepared for the BGS Internet.

The developing groundwater drought has increased media interest in groundwater level data and the Hydrological Summaries, and an animated timeline of groundwater level will be released on the BGS website.

FY 2012-13

The drought will almost certainly dominate activities, with a requirement for more regular loading of data in drought prone areas.

Ongoing discussions with EA about acquiring all their logged and telemetered data for the archive, rather than the small subset currently managed, will continue.

4.4.3 Hydrogeological data

FY 2011-12: The groundwater fauna database was developed further, and is being used to record groundwater biodiversity at national level.

A joint CEH/BGS Pathway project on groundwater surface water interaction led to development of a shared model for storing project data with CEH, that should see further expansion of capacity to ingest logged and telemetered data to corporate databases.

FY 2012-13: Further population of ecological and surface water interaction databases will be required.

Planned work on looking at how climate change related datasets are stored will continue.

4.4.4 Other Tasks in IM3

FY 2011-12: Commercial funding has supported work on hydrogeological database design in India.

Work on a digital Hydrogeological map has significantly improved detailed metadata on the maps BGS holds, and provided a core spatial dataset for WellMaster data conditioning.

4.4.5 Future Tasks 2012-13

Significant paper record holdings for the UK and overseas need ongoing management, rationalisation and scanning. The success of the South African Development Community (SADC) 'Grey Literature Database' (<http://www.bgs.ac.uk/sadc/index.cfm>) may provide a model for other regional datasets and some of the BGS UK hydrogeological information.

4.5 INFORMATION MANAGEMENT IN GEOCHEMISTRY 2011-12 AND BEYOND (IM4)

Susan Hobbs

4.5.1 Laboratory Information Management System (LIMS)

A new LIMS has been purchased and installed to replace the previous LIMS which was considered to be inflexible and complicated. All data has now been transferred from the old LIMS, which was on a standalone sever, to the new LIMS which is on the SAN. The old LIMS has now been decommissioned. The remaining rationalisation and restructuring of the new W:\SciFac\AG SAN space is nearing completion.

4.5.2 G-BASE Audit

Data from the London Earth Project including urban soils collected in 2005 and 2006 which lie outside the Greater London Authority boundary, has been conditioned (with respect to Certified

Reference Materials) and loaded to the geochemistry database. 7,189 samples were taken and analysed for 60 analytes (431,340 data points/records).

The G-BASE 2005 and 2006 Regional Soils were reconditioned with respect to Certified Reference Materials – 2,235 records for 58 elements.

Analytical data for the Regional 2007 surface soils (48,836 records) and the Regional 2008 surface soils (34,220 records) have been loaded.

Field data for Clyde Basin has been loaded (1,230 records).

Previously missing samples from the London Earth data set have been recovered, analysed, conditioned and are now available (1,160 records). All available legacy G-BASE raw analytical data has now been brought together into one folder on the W: drive, dating back to the Borders Atlas, as part of the back population of the Raw data tables.

4.5.3 Water samples

The majority of the 2010 water samples have been analysed and audited, however as a result of the transfer of the XRF laboratories from BGS to PanAnalytical and the resulting physical office, laboratory moves and refurbishment, there has been a delay in reporting some of the data. There are 7 anions and organic carbon data still outstanding. This delay has resulted in staff no longer being available to completely finish the audit this financial year.

4.5.4 Geochemistry Database changes

The Geochemistry database has been amended to include a sample processing field. For some methods of analysis, this is as important as the method of analysis. A new preparation table and dictionary have been created to address this issue.

Two new method codes have been added to the analytical method dictionary to accommodate measured and estimated inorganic carbon data. 5,777 measured and estimated inorganic carbons have now been added to the database.

4.5.5 Future Tasks 2012-13 – LIMS

Data was systematically loaded directly into the Raw data tables from the LIMS, however with the move of the XRFS laboratory to PanAnalytical, placing them outside the BGS “fire wall” this process no longer takes place. Code is to be written to enable this process of systematic loading of data (now received from PanAnalytical) to continue, allowing the outstanding raw data for 2011 to be loaded.

4.5.6 Future Tasks 2012-13 - G-BASE

It is planned to continue with auditing and capturing new and legacy data. There may be an increase in the number of samples collected over the next few years as a result of the winding down of the G-BASE project. The 2012 field data should cover Cornwall and Devon, possibly Somerset too. In order to complete the coverage of the whole of the UK over the next few years, the number of samples collected per year may have to increase.

Regional and urban Clyde Basin soils are to be conditioned.

Data conditioning of 2006, 2007, 2008 and 2009 stream sediments is in progress.

4.5.7 Future Tasks 2012-13 - Water samples

Data for the 2010 water samples was received in May 2012 and the audit of the data is due to be completed this financial year. As a result of the “winding down” of G-BASE the coverage of the sampling this year is likely to increase and potentially the samples from SW England area will be collected.

4.5.8 Future Tasks 2012-13 – Radon

The Radon data is being reviewed and will be stored sorted into a single folder. The decision as to whether the radon data should be archived or stored on the S drive will not be made until the review and centralisation of the data is complete, allowing the extent of the complexity to be ascertained.

4.6 COASTAL AND MARINE INFORMATION MANAGEMENT (IM6) 2011-12 AND BEYOND

Paul Henni

4.6.1 BGS Marine Data Management Standards for Offshore Surveys

The BGS Marine Data Management Handbook (Open Report, Crummy 2008) has proved to be successful in standardising BGS Marine Survey Data Acquisition. Data delivered post-survey to the onshore Data Management Team is now received in a standard way, streamlining processes for databasing and storage.

Internally, elements of the handbook have been revised, together with updates to our Data Management procedure documentation and quality control processes. This includes Multibeam data where there are now clear ways of working and storing data based on recommendations from the BGS Hydrographic Survey Manager.

An updated version of the handbook will be produced after review of new MEDIN Seabed Survey Guidelines.

The Marine Data Management Team have worked closely with those managing the SAN and the BGS Archive systems to ensure the Marine data is stored to corporate standards.

4.6.2 Accessioning (incoming data and information)

The accessioning system ('data post-room') has been publicised amongst the Marine Geologists and any significant data incoming to BGS is generally booked in prior to storage in an appropriate area.

4.6.3 The Coastal and Marine Database (CMD)

Major work has been done to add to this, the heart of the Data Management system:

The Offshore Reports collection now has a completely revised design incorporating appropriate corporate dictionaries and is in the final stages of being released to the Corporate schema. The Team have worked closely with the NGRC and undertook a 'stockcheck' to ascertain availability of reports.

Geological sample information has been extended to allow the storage of parametric and other 'extended' data and work is ongoing to QC and migrate legacy digital and paper data.

Multibeam data has also now been incorporated into the design.

Other maintenance and enhancement work has also been done to allow external delivery of our data compliant with external standards and requirements.

This work was undertaken primarily by IM6 staff, working closely with the BGS Data Architect. IM6 have also liaised with IM2 staff to ensure the old onshore/offshore data boundary is further eroded.

4.6.4 Creation of Metadata to Meet Required Standards (INSPIRE, MEDIN, GEO-SEAS)

IM6, working with the web systems team, have provided MEDIN format metadata to the MEDIN Data Portal.

Geo-Seas metadata has been created and geological sample data released.

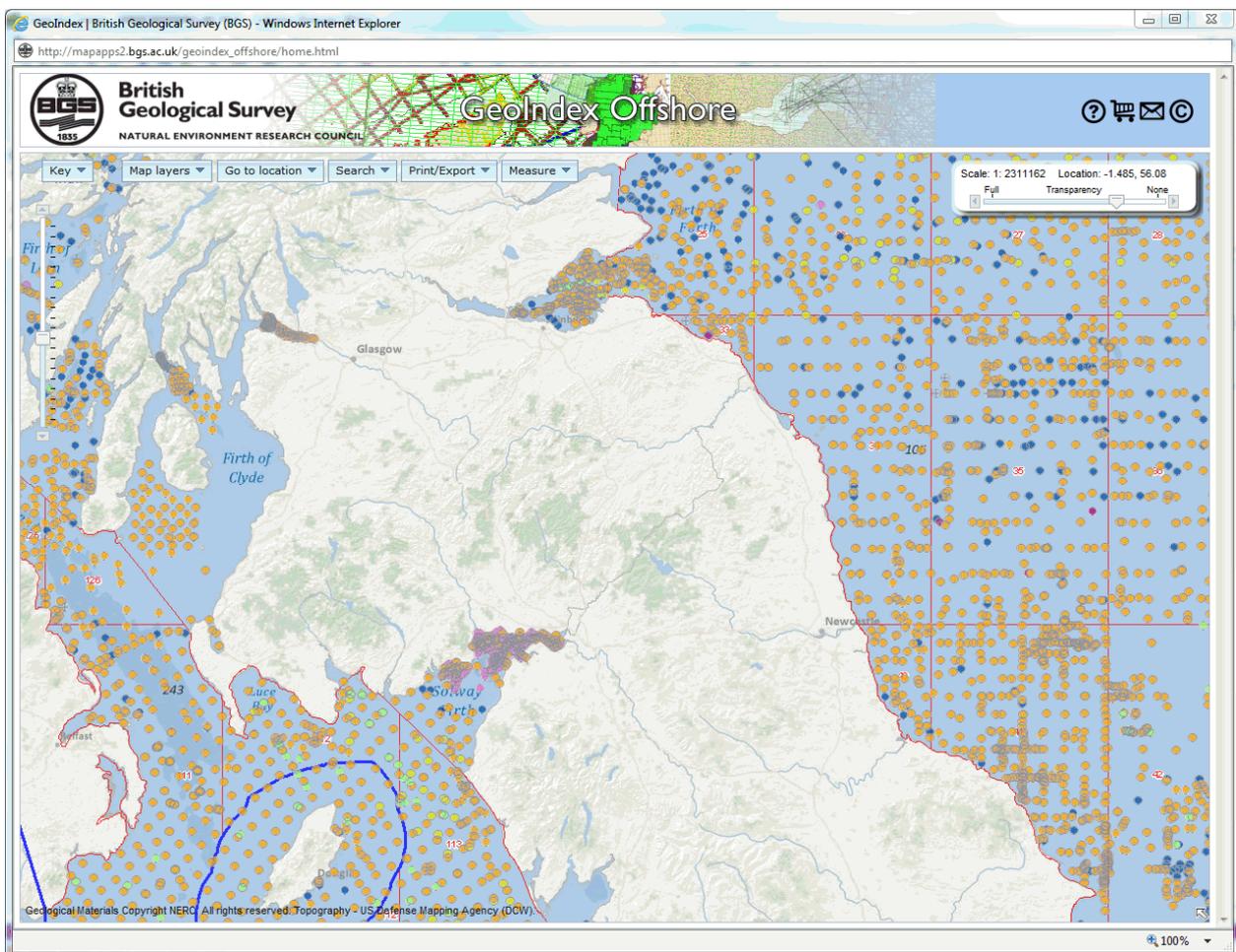
A long overdue Marine Metadata Review was undertaken to ensure that BGS meets requirements for MEDIN and also INSPIRE / UKLP. It is hoped that this will be implemented soon (at time of writing).

4.6.5 Delivery of Data and Information

Progress has been slow on this due to ongoing internal debates about the status of Marine Data; whether or not it can be delivered as open data. Discussions on technical issues and requirements for setting terms and conditions continue.

Currently, IM6 are working to deliver 'Open' geological sample data via both the BGS Offshore Geoindex and to Geo-Seas. Once the model is working, it will be extended to other data types.

Figure 3 Publically available BGS Sample Data locations



The Offshore Geoindex will be used as the delivery mechanism for Marine Geology and Geophysics Open Data.

Requests for BGS marine data are received as email generated from the Offshore Geoindex or through central enquires and are dealt with by the small Marine Enquiries Team. The ultimate outcome of current developments will be self service for data customers.

4.6.6 MEDIN

The BGS/NGDC will continue to operate as a MEDIN Data Archive Centre (DAC) and maintain active participation in MEDIN (including in-kind BGS contribution from IM6 funds).

IM6 have had a successful year as a MEDIN DAC and participant. BGS have regularly attended workshops, working groups and the MEDIN Open Meeting contributing to the ongoing goals and tasks.

17 MALSF Survey Datasets were successfully archived, a major undertaking given the age and organisational state of some of them. This was achieved by several IM6 staff despite competing demands for their time.

The MEDIN Oil and Gas Site survey work continued, with the database model migrated to the corporate schema, completion of the loading of legacy survey information and the ongoing capture and delivery of information about new surveys.

A test set of raw Multibeam Backscatter data has been received and BGS will continue to receive and manage these data to build a full dataset. This work will be a BGS contribution to MEDIN and will make this set of data available for BGS use.

Work is ongoing on the MEDIN funded archiving of JNCC and Wessex Archaeology survey datasets.

4.6.7 BGS or BGS/Partners marine surveys

Due to a combination of MEDIN work and other priority tasks, progress on processing new and immediate legacy (post-2000) BGS or combined BGS/Partners marine surveys has slowed. Additionally, the retirement of three members of the Marine Team has had most effect in this area of work.

It is hoped that both the resources and staff can be made available by BGS to allow this work (and more) to continue.

4.6.8 Data Rescue

Ongoing capture of paper seismic records continues, led by customer enquiries and BGS science needs.

Several areas of interest, primarily in the North Sea, were captured.

Additionally, work undertaken for the BGS Crown Estate Aggregates Project highlighted legacy paper records with potential data for capture and existing database tables requiring significant work to review, QC and migrate. This work is ongoing and will go to improve BGS seabed sediment maps.

4.6.9 United Kingdom Hydrographic Office (UKHO) Data and Information 2011-12

The Marine Data Management Team worked with the NGRC team to integrate, index, manage and archive disparate paper and digital United Kingdom Hydrographic Office (UKHO) datasets. These are either (a) already held, requiring review and potential culling of obsolete (primarily paper) records, or (b) new to BGS for receipt as part of an MoU between BGS and the UKHO. These datasets, especially the newer digital ones, are required to underpin the Offshore Science Programme of work.

The UKHO paper record collection at Loanhead was visited prior to the retirement of the staff member with the greatest corporate knowledge of these data, to ensure the passing on of his understanding to NGRC colleagues.

On the digital side, IM6 have received and are storing UKHO/MCA multibeam data for BGS use.

4.6.10 Other ongoing work planned for 2011-2012 in IM6

It was planned to further develop the Marine Enquiry procedures to facilitate transfer of routine enquiries across to the BGS Central Enquiries Team. This has been put back to 2012-13.

The passing of the management of analogue data and paper records to NGRC colleagues, while providing advice and knowledge transfer, is to be continued.

The development of the Marine Data Management Team and subsequent knowledge sharing is to continue.

There were plans to work with the Landmark Geophysical Data Manager to improve knowledge and data management of this system and data, including legacy tape issues. This work was limited by staff availability.

Completion of the definitive set of BGS-held BIRPS data and information was planned (subject to staff resources).

BGS data and information holdings were to be publicised widely.

4.6.11 Desired Tasks 2011-2012, subject to funding

Other desired tasks which were generally not possible to fund and were worked on in a limited way only:

Completion of the review and integration of the disparate 'Marine' and 'Coastal' analogue and digital collections into a unified and coherent indexed and delivered set of information (working across all BGS areas of work).

Management of legacy projects and material received from retired colleagues.

Management of legacy survey data and information pre-2000, plus non-BGS survey data and information (outwith MEDIN DAC data).

4.6.12 Other Future Tasks 2012-13

Review all tasks and re-prioritise as necessary, e.g.

Following feedback from younger geologists, there is often a lack of knowledge of previous project results. There is a need to re-visit the suspended project data harvesting work to dig out the nuggets of useful data and interpretations and make these easily available.

The problems found with legacy PSA data have required a large amount of work to improve quality. Work is now in progress to create unified and QC'd data plus desired outputs.

There are now complex linkages amongst Sample data, Offshore reports and Surveys. Continued 'detective work' is important in identifying data sources, ownership and status (open/restricted), to allow easy understanding of whether data can be used by BGS or made more widely available as 'Open Data'.

Building on the Offshore Reports work and the same requirement to know the status of reports held by BGS, IM6 will work with NGRC on a campaign to contact and agree report status with IPR owners.

Once the model is working for delivering 'Open' geological sample data via both the BGS Offshore Geoindex and to Geo-seas, it will be extended to other data types.

Loss of three team members has had a negative effect on progress. There is potential difficulty in adding to the team due to the specialist knowledge required for some tasks, but discussion is ongoing with Skills Leaders to resolve this.

A Team task list with more detailed information on activities is stored within the project.

5 European Co-Funded Activities

5.1 GEO-SEAS

Helen Graves

Geo-Seas, a Pan-European e-infrastructure project for the management of marine geological and geophysical data, has the aims:

1. Creation of a unified European e-infrastructure for marine geoscientific data.
2. Facilitation of the locating, accessing and delivery of federated marine geological and geophysical data sets.
3. Provision of user access to harmonised marine geological and geophysical metadata and data via the Geo-Seas portal.

Work has continued on the development of interoperability with other data types within other disciplines. The annual progress report to the European Commission was accepted and approved allowing continuation of the project into 2012 – 2013.

The table below is a list of the outputs from the Geo-Seas project for 2011/2012. Key outputs are highlighted. The most important are the user workshop at the Oceanology International exhibition where Geo-Seas hosted a public workshop during which the tools and services being developed for Geo-Seas were showcased to the end users for the purposes of raising awareness and also getting feedback for the final release of the tools.

In addition to the Geo-Seas presentation at the EGU meeting in Vienna, Geo-Seas also hosted a splinter group meeting on marine data management which included representatives from Europe, USA and Australia. The purpose of this meeting was to discuss potential collaborative opportunities with the objective of furthering the development of a common approach to marine data management on a global scale.

Geo-Seas also co-authored an article printed in the December 2011 edition of the International Innovations magazine entitled: Uniting marine data management.

Table 4 Geo-Seas Outputs 2011 - 2012

Date/Venue	Title/Concertation Activity	Type	Author* (Partner)
3-6 May 2011, Helsinki, Finland	GeoHab 2011 international conference: “Standardisation and Harmonisation in seabed-habitat mapping: How can a geological data-infrastructure project contribute ”	Poster	Vera van Lanker (MUMM) on behalf of the Geo-Seas consortium
28 June - 1 July 2011, Brest, France	The Future of the 21 st century ocean: Marine Sciences and European Research Infrastructures - an international symposium: <i>Geo-Seas</i>	Poster	Helen Graves (BGS) , Dick Schaap (MARIS)

Date/Venue	Title/Concertation Activity	Type	Author* (Partner)
2 September 2011	Geological data accessibility in the European framework. International Symposium “Human Footprint on the Seafloor. Keys from the Past. Doors to the Future”. Brussels, Natural History Museum (BE),	Oral	Vera van Lancker (MUMM) , K. De Cauwer, C. Baeteman, S. Scory, (RBINS-MUMM)
Brussels, Natural History Museum (BE), 2/9/2011	International Symposium “Human Footprint on the Seafloor. Keys from the Past. Doors to the Future”. A digital seabed-sediment database for northwestern Europe: flexible query possibilities for multiple end users.	Poster	S. van Heteren (TNO), V. Van Lancker, M. Zeiler, J. Leth, J. Valerius, P. Kiden..
14 September 2011; Brussels - Belgium	EU-US cooperation workshop on research infrastructure interoperability across the Atlantic	Oral	Dick Schaap (MARIS)
28 – 29 September 2011, Santa Barbara, California, USA	Environmental Information Conference: <i>Geo-Seas – a pan European marine geoscientific e-infrastructure</i>	Oral	Helen Glaves (BGS)
September 2011	EPOS Newsletter: <i>Geo-Seas – a pan-European infrastructure for the management of marine and ocean geological and geophysical data</i>	Newsletter	Helen Glaves (BGS), Dick Schaap (MARIS) Gabriela Carrara (LNEG) and the Geo-Seas consortium partners
September 2011 Brussels, Belgium	Quest 4D Final Conference: A digital seabed-sediment database for northwestern Europe: flexible query possibilities for multiple end users.	Poster	S. van Heteren, S., V. Van Lancker, M. Zeiler, J. Leth, J. Valerius, P. Kiden. (TNO)
October 2011, Zadar, Croatia	SPLASHCOS project meeting	Oral	Sytze Van Heteren (TNO)
9 – 12 October 2011, Minneapolis, Minnesota, USA	Geological Society of America annual meeting: <i>Geo-Seas – a pan European marine geoscientific e-infrastructure</i>	Oral	Helen Glaves (BGS)

Date/Venue	Title/Concertation Activity	Type	Author* (Partner)
October 2011	Studiedag Mariene Aggregaatextractie: noden, richtlijnen en toekomstperspectieven. Geological data accessibility in the European framework.	Oral	V Van Lancker K. De Cauwer, C.Baeteman, S.Scory, (RBINS-MUMM).
16 November 2011, Istanbul, Turkey	GEO VIII Plenary meeting: <i>Geo-Seas – a pan-European infrastructure for the management of marine geoscience data</i>	Oral	Helen Glaves (BGS)
16 -17 November 2011, Istanbul, Turkey	GEO VIII (Group on Earth Observations Plenary) EU stand	Poster	Geo-Seas partners
November 2011, Tallinn, Estonia	EMODNET Geology meeting: <i>Geo-Seas – a pan-European infrastructure for the management of marine and ocean geological and geophysical data</i>	Oral	Helen Glaves (BGS)
7 December 2011, Amsterdam, Netherlands	Hydrographic Society Benelux Workshop Geophysics: Practice, Theory and Data Management?: <i>Data exchange. The future of geophysical and geological data management in EU frameworks</i>	Oral	Cees Laban (Geo-Seas Advisory Board)
9 December 2011 Sofia, Bulgaria	Geosciences 2011. Annual Conference of the Bulgarian Geological Union	Poster	Lyubo Dimitrov and the Geo-Seas partners
24 February 2012 Bruges, Belgium	VLIZ Young Scientists Day. Marine scientists meeting: Geological data accessibility in the European framework.	Poster	V. Van Lancker, K. De Cauwer, C.Baeteman, S.Scory, (RBINS-MUMM)
24 February 2012 Bruges, Belgium	VLIZ Young Scientists Day, Marine scientists meeting: Data services at the Belgian Marine Data centre: from event logging onboard, historical data archival to standardized geological data.	Oral	K De Cauwer, M. Devolder, R. Lagring, S. Scory and V. Van Lancker (RBINS-MUMM)

Date/Venue	Title/Concertation Activity	Type	Author* (Partner)
24 February 2012 Bruges, Belgium	VLIZ Young Scientists Day, 02/2012. Marine scientists meeting: Demonstration of geological data products. Seabed viewer. Belgium (300+ participants)	Software demonstration	V Van Lancker, K. De Cauwer, (RBINS-MUMM)
March 2012	Oceans and Lakes programme University of Gent, University of Brussels & University of Antwerp: In-situ and remote sensing tools in aquatic environments. Lectures in which Geo-Seas is promoted to students.	Oral	V. Van Lancker , (RBINS-MUMM)
14 March 2012 London, UK	Oceanology International 2012 exhibition: Geo-Seas user workshop	Oral	Geo-Seas partners
April 2012 Oostende, Belgium	2012. NATO Training course: Multibeam and side-scan sonar: an introduction.	Oral	V. Van Lancker (RBINS-MUMM)
23- 27 April 2012 Vienna, Austria	European Geosciences Union General Assembly : Geo-Seas: a pan-European e-infrastructure for marine geoscience data management	Oral	Helen Glaves (BGS)
23- 27 April 2012 Vienna, Austria	European Geosciences Union General Assembly: Collaborative work and tools towards wide scientific community driven metadata model and vocabulary building. the case of the EU Geo-Seas Project.	Oral	Paolo Diviacco (OGS) , Roy Lowry (BODC) and Adam Leadbetter (BODC)
23- 27 April 2012 Vienna, Austria	EMODNET: what next?	Oral	Dick Schaap (MARIS)
23-27 April 2012	Maine Data Management Splinter Group meeting	Oral	Helen, Glaves (BGS), Dick Schaap (MARIS)

Date/Venue	Title/Concertation Activity	Type	Author* (Partner)
2011	<u>Marine Seismic Metadata for an Integrated European Scale Data Infrastructure. The Fp7 Geo-Seas Project, BGTA in print, DOI: 10.4430/bgta0051</u>	Paper	Diviacco, Paolo, Roy Lowry and Dick Schaap
2011	Research of the bedrock geology of the Central Baltic Sea. Baltica. Vol. 24, No. 1, p. 1–12.	Paper	Algimantas Grigelis (NRC-IGG)
2011	History and problems of geological research of the south-eastern Baltic Sea. Baltica. Vol. 24, Spec. Iss.: Geosciences in Lithuania: challenges and perspectives, p. 7–12.	Paper	Leonora.Ž Gelumauskaitė (NRC-IGG).

6 Additional Tasks

6.1 STRATEGIC DELIVEABLES

The main aim of this task is to ensure that project information assets currently lying inactive and unused on the SAN are recovered, made ready for use and re-use and incorporated into corporate records where they form the basis of products and services. In short - recovering assets from completed projects and conditioning them for further integrated use.

Research data are a valuable resource, often requiring much time and money to be produced. Many datasets have a significant value beyond the original research. These outputs should not be “gathering dust” but should be actively managed. Data must meet quality thresholds and be stored in a reliable manner supporting their promotion and reuse.

6.2 MARINE DATA MANAGEMENT (PAUL HENNI)

Marine Survey Data are a valuable resource; there is a continuing need to apply relevant quality control procedures to existing marine survey data before loading it into the corporate Marine Survey database.

Significant amounts of data have been generated by a large number of Marine Surveys and there is considerable work to be done to ensure the quality of these. This Survey work is being broken down into achievable work packages to ensure steady, measurable progress. The longer the data is left unprocessed, the more time will be required to retrospectively do the work, so it is important to maintain the current momentum. All of this work feeds in to the BGS role as a MEDIN Data Archive Centre (DAC).

The result of this work is enhanced, quality controlled Marine Geoscience data, stored to corporate standards, with appropriate elements stored in the BGS offline archive system. This enables improved delivery of revised datasets to internal and external customers.

Specific deliverables include:

- Successful delivery of MEDIN-compliant metadata to the MEDIN web portal and updated BGS Offshore Geoindex.
- Completed review and update of post-2000 Marine Surveys. Initiate work on legacy surveys (pre-2000).
- Up to date quality controlled Marine Sample information (clear 5 year backlog of data loading).
- Support to the Marine Mapping Programme: completed scanning, archiving and vectorisation to SEG-Y of seismic records (for appropriate data types) of target offshore regions, e.g. Forth Approaches, and others to be agreed with Marine Team Leaders.

6.3 3D FOSSIL SCANNING (TIM MCCORMICK)

BGS Collections' recent purchase of a 'NextEngine' 3D scanner, together with the availability of Assetbank for storage and web delivery of digital assets, provides the opportunity to revolutionise access to our palaeontological collections. Scientists and the public will be able to view, download and manipulate 3D digital representations of our specimens. This work was to carry out a pilot project to scan selected specimens, representing a range of geological ages, biological phyla, and sizes, and create a short 3-5 minute 'Digital Object Presentation'.

The 3-5 minute digital presentation was displayed in the 'digital object' session at the World Conference on Marine Biodiversity, Aberdeen, 26-30 September 2011 and is available on the BGS website.

Specific deliverables included:

- 3D scans of a range of fossils selected from the BGS palaeontological collections.
- A written best practice procedure for generating and storing 3D scans of fossils and associated metadata, to guide future 3D scanning activities.

The proposal was strongly supported by the Head of Information Management, impressed by this new way of presenting fossils believing it to have the potential to make the collections much more accessible for research, education and the public; the technology becoming affordable at a bad time in the economic cycle.

This work was able to produce a number of examples to demonstrate the capabilities of the new system and contributed to the correct environment for obtaining the JISC grant mentioned in section 4.3.1.

6.4 MARINE MINERALS (TOM BIDE)

Capture of currently missing PSA data into a digital format compatible with BGS corporate marine database.

This task was to capture data; currently only held in disparate paper based sources and appearing as gaps in coverage on BGS datasets, for inclusion in the BGS corporate Oracle marine database which underpins numerous marine data products and External Income work, critically the current offshore minerals project. Specifically the work was to capture particle size analysis data from sample sheets, map plots and other relevant sources, for inclusion in BGS corporate databases for areas where data is currently missing.

The aim is to greatly enhance BGS corporate datasets by producing a dataset with complete coverage of the UKCS, useful to BGS projects and commercial work where a national dataset of

particle size analysis is required, such as that being currently undertaken on behalf of the Crown Estate by the Mineral Resources and Policy Team.

Specific deliverable:

A database containing currently non-digital particle size analysis data adhering to BGS corporate standards.

All identified data that was available to BGS in paper form but not digitally captured for England and Wales has been digitised. Digitised particle size analysis data has been stored in excel documents and has been passed to data managers of marine data in Edinburgh for inclusion into a final corporate database, which is due to be complete by the end of 2012, as part of a wider review into BGS's seabed sediment sample data holdings. All undigitised mapping data from sediment thickness maps and bed form maps present on published BGS 2:250 000 marine geology maps has been captured for England and Wales and data has been passed to managers of marine spatial data in Edinburgh for appropriate storage and use.

6.5 EARTH OBSERVATION FRAMEWORK (EOF) (ROB ARMSTRONG)

This task is involved with keeping the BGS contribution to the Earth Observation Framework up to date and relevant and is under the direction of Andrew Howard.

The UK Environmental Observations Framework (now subsumed into LWEC) maintains a catalogue of environmental observation activities carried out across government in the UK. BGS and other research centres are asked to contribute updates to their entries in that catalogue on an approximate annual basis.

The requirement this year was to provide the spatial metadata; with no requirement to update the whole catalogue of BGS activities. In previous years the UK-EOF catalogue had not gathered any systematic spatial information on these activities. This year, they were collecting spatial metadata (i.e. the grid references) for monitoring sites and observatories. These cover locations where continuous or periodic measurements or observations are made but exclude baseline survey, mapping and sampling activities. The exercise therefore was relevant to only a subset of the BGS activities included in the UK-EOF catalogue.

The UK-EOF Management Group had agreed that an exercise to map public funded observations sites in the UK this year would be a useful exercise and Members were committed to providing site information. The site data will be used by Natural England (NE) to construct a collaborative map. They have already carried out co-incidence mapping with 22 observations networks across the UK and will now work to extend this to include UK-EOF Management Group organisations information. To quote:

This site information will;

- 1. create potential collaborative opportunities for UK-EOF Members and*
- 2. assist NE by increasing the co-incidence mapping of long term observation sites to meet certain internal aims (with a specific focus on atmospheric pollution).*

The request was for site information to be sent in Arc GIS shape files or Grid references (in a spreadsheet) along with the relevant UK-EOF record (downloaded from the catalogue) to provide additional information about the activity being mapped.

Rob Armstrong completed the work and provided:

- Locations of BGS observation sites in ESRI shape file format
- An excel table cross-referencing the sites to the UK-EOF catalogue (by catalogue number). This table indicates which BGS activities we have included/excluded, and why (essentially,

we have included sites where we carry out continuous monitoring or regular periodic observations, and have excluded sites of one-off baseline observations, and sites where we are not the lead organisation).

- A pdf map of UK showing the site locations, for reference.

6.6 OTHER ACTIVITIES

- Archiving CA Seismic and Well-log Tape Data: The departure of a key member of staff made it imperative to transfer knowledge and skills in this area to a successor who could continue with the work.
- Loading of legacy G-BASE stream sediment/soil raw data results to the raw data tables in the Geochemistry Database. The data from Borders to Humber-Trent was necessary as control samples and raw data until 2005 were not captured into a corporate database. (G-BASE has always loaded processed data into the main Geochemistry Database analyte determinations table).

7 Conclusion

The IMFO Team is engaged on a wide range of data activities which underpin the scientific research and commercial activities of the BGS. The work is essential for the long-term security and validity of the most fundamental and basic data central to the effective working and reputation of the organisation.

Research data are an institutional asset requiring stewardship. Virtually all aspects of BGS scientific work depend upon the routine use of reliable and up-to-date data and information. Effective exploitation by many systems and users requires effective data management to maximise initial investment made in creating or gathering data. Reliable re-use of data is only possible if materials are curated in such a way that their authenticity and integrity are retained.

BGS has a high investment in research. Data can be very expensive to capture and analyse and often impossible to recreate once lost. Observational data (by definition) is irreplaceable. Current generations of instruments can gather more data than can be analysed. There is an established need to rework data as new methods become available; secure storage of well defined, findable, accessible data is important for the science base of the country.

The NERC Science Information Strategy has released its guidance documentation on Data Management Planning and produced a Data Value Checklist for research data accessioning by the data centres. NERC now has a mechanism for the minting of Digital Object Identifiers (DOIs). Datasets can be cited and will feature in organisational metrics together with research papers. Many journals now require links to the underlying data that support research papers. This is the age of Open Data, there is a mandate to ensure that government funded research outputs are accessible to all, the work of the IMFO Team has never been more important or more necessary.

Appendix 1 INSPIRE

Scope of the INSPIRE Directive

This Directive shall cover identifiable collections of spatial data, hereinafter “spatial data sets”, which fulfil the following conditions:

- (a) they are related to an area under the jurisdiction of a Member State or to its exclusive economic zone/search and rescue region, or equivalent;
- (b) they are in electronic format;
- (c) they are in the possession of either of the following:
 - (i) a public authority, having been produced or received by a public authority, or being managed or updated by that authority;
 - (ii) a natural or legal person on behalf of a public authority;
 - (iii) a third party to whom upload services have been made available in accordance with Article 17(3);
- (d) they relate to one or more of the themes listed in Annexes I, II or III.

<p>Annex I Coordinate Reference Systems Geographical Grid Systems Geographical Names Hydrography</p>	<p>Annex II Elevation Land cover Orthoimagery Geology</p>
---	--

Appendix 2 Data Holdings Maintained by IMFO

This list comprises the main databases and datasets maintained by the IMFO Team. The list includes the significant corporate datasets, other smaller, more specialist data holdings, also maintained by IMFO, do not feature here.

Dataset	Description
Borehole Geology	Database of geological interpretations of boreholes drilled in Great Britain. These interpretations may have been supplied by the geotechnical company responsible for drilling the hole, or may have been made subsequently either by BGS geologists or other agents. Borehole Geology is heavily used in conjunction with the Single Onshore Borehole Index (SOBI) to generate 3-D models and other representations of the geology of Great Britain. Currently just over 3 million records.
BritPits	Also known as the Mines and Quarries Database, this stores information about active and inactive onshore mineral workings, also rail depots and wharves used in movement of commodities. Compiled and used by BGS to periodically produce the 'Directory of Mines and Quarries' for sale through the MineralsUK website, and for market intelligence / analysis and resource planning by internal and external customers including local councils. Data made externally available under licence. Currently approximately 140,000 records.
Coastal / Marine Database (CMD)	Database of BGS-held Marine Survey data, including an index at survey level linked to digital and analogue geophysical, sample and multibeam data and information.
Data Management Planning System (DMPS)	The system where the information on project data management is entered and stored. It is a tool for data management planning but also for capturing long-term data assets and for management of the Storage Area Network resource.
Discovery Metadata	Database holding ISO 19115 compliant metadata describing significant data holdings in BGS.
Geochemistry Database	Database for geochemical data from the land area of Great Britain and Northern Ireland (excluding groundwater). The Geochemical Baseline Survey of the Environment (G-BASE) is the most active contributor, about 60%. About 35% samples are from the Minerals Reconnaissance Programme (MRP). The remaining data is from various other BGS projects. Samples types include stream sediments, soils, stream waters, rocks, drill cores, panned concentrates and tills. There are now over 10 million analyte determinations stored in the database.
Geoscience Imagebase 2 / Geoscentic	This is the BGS database of digital geoscientific images and associated metadata, and includes images taken by BGS geoscientists as well as collections accessioned from other organisations. A subset of the images are selected and published on the BGS website as 'Geoscentic'.

Dataset	Description
Issues Log	Database and tracking system for issues identified as possible errors in corporately held data.
Lexicon of Named Rock Units	Database of definitions of rock units and other geological deposits recognised by BGS on the UK and its continental shelf. As well as being a repository of knowledge, the Lexicon is used to constrain the lithostratigraphical attribution of many other BGS databases. It is searchable on the BGS website. Currently approximately 13,200 records.
National and Global Geomagnetic database	Geomagnetic data from BGS observatories in the UK and overseas.
National and Global Seismic database	<p>Continuous seismic waveform data for the British national seismic network for up to 12 years.</p> <p>Parametric data for national and global seismic events going back over 1000 years.</p>
National Building Stones Database	Database of (mostly historic) buildings and construction materials. The information are collected and supplied by English Heritage staff. BGS uses the information to answer enquiries and provide advice to customers and partners, often local authorities and heritage agencies. Approximately 4000 index records plus additional related data.
National Geotechnical Properties Database	Database of geotechnical properties and other information derived from site investigation reports supplied to BGS. BGS compiles this information for use in engineering geology research, answering enquiries, and increasingly to attribute 3-D geological models. Currently approximately 81,500 index records plus additional related data.
National Groundwater Level Archive	Database of groundwater level observations from representative boreholes across the British Isles. The majority of data are collected by environmental regulators or reported by water well owners, but increasingly data from research networks is being included. Working closely with the National River Flow Archive, which is operated by the Centre for Ecology and Hydrology's national river Flow Archive, the data are used to produce monthly and annual Hydrological Summaries, and are a key data source for assessment of water resources and for analysing long term hydrological trends. Currently holds up to date time series for around 200 boreholes, historical time series for a further 1500 sites with around 50,000 sites having intermittent measurements.
National Well Record Archive/WellMaster	Database of groundwater data, compiled from the paper well record holdings of the National Well Record Archive, which receives copies of all water wells drilled in the UK, under provisions within the Water Resources Act. The digital database, WellMaster, holds geological and hydrogeological data used for hydrogeological research, answering enquiries and product development. Ancillary databases hold detailed aquifer property and groundwater chemistry data. Currently approximately 110,000 records are held, although the level of detail available is constrained by original record quality.

Dataset	Description
World and UK Mineral Statistics	Database of world mineral production statistics. Compiled and used by BGS for market intelligence / analysis, and for periodical production of 'World Mineral Production', 'European Mineral Statistics', and 'United Kingdom Minerals Yearbook' for sale or download through the MineralsUK website. Currently approximately index 241,000 records plus additional related data.

Glossary

AGI – Association for Geographic Information

BIRPS – British Institutions Reflection Profiling Syndicate

BS ISO 19115:2003 – The British Standard for Geographic information – Metadata.

CEH – the NERC research centre for Ecology and Hydrology

CMD – Coastal and Marine Database

DAC – Data Archive Centre [MEDIN]

DEFRA – Department for Environment, Food and Rural Affairs.

DMPS – Data Management Planning System, an intranet application for recording details of data management activities for a particular project, service or dataset. Information is recorded in Oracle and is used for data asset and SAN management.

EA – Environment Agency

EGU – European Geosciences Union

G-BASE – Geochemical Baseline Survey of the Environment

Geo-Seas - An e-infrastructure project of 26 marine geological and geophysical data centres, funded under EU FP7.

GI2 – Geoscience Imagebase 2, the BGS database of geoscientific images, which is hosted and served using Assetbank software.

INSPIRE – A European directive aiming to make consistent information on spatial datasets about the environment available and to create services for accessing these datasets.

IPR – Intellectual Property Rights

IUGS- CGI – International Union of Geological Sciences - Commission for (the management and application of) Geoscience Information.

JNCC – Joint Nature Conservation Committee, a statutory adviser to UK Government and devolved administrations

LIMS – Laboratory Information Management System, a system designed to manage data output from laboratory equipment and activities

LUN – Logical Unit, virtual disk of the Storage Area Network, comprised of groups of physical disks.

MCA – Maritime and Coastguard Agency

MEDIN – Marine Environmental Data and Information Network

Metadata – Data to describe data, e.g. Discovery Metadata, data that describes a dataset so that other people can decide if it's useful to them, or Technical Metadata, data that describes database objects and their relationships.

MoU – Memorandum of Understanding.

MTG – The Multilingual Thesaurus of Geoscience, a working group under IUGS-CGI.

NERC Data Catalogue Service – A service to allow users to search a catalogue of metadata (information describing data) to discover and gain access to NERC's data holdings and information products. The metadata are prepared to a common NERC Metadata Standard and are provided to the catalogue by the Data Centres.

NGDC - National Geoscience Data Centre (includes both records and collections)

NGRC – National Geoscience Records Centre

OpenGeoScience - A free service on the BGS website where the public can view maps, download photographs and other information.

PSA – Particle Size Analysis

RMS - Resource Management System, a NERC wide resource for project initiation and planning available via the Intranet.

QC – Quality Control

SAN – Storage Area Network – a high-speed network of various levels of storage hierarchy that permit large volumes of data to be stored in a secure manner.

UK GEMINI - (Geo-spatial Metadata Interoperability Initiative). UK GEMINI is a specification for a set of metadata elements for describing geospatial data resources for discovery purposes. It has been produced and is maintained by the AGI Standards Committee as a free of charge resource.

UKHO - United Kingdom Hydrographic Office

UKLP – UK Location Programme. Formed to help and advise organisations to implement the UK Location Strategy and INSPIRE. The UK Location Programme team works closely with the devolved administrations of Wales, Scotland and Northern Ireland.

XRF – X-Ray Fluorescence, a technique used by BGS for geochemical analysis.

References

British Geological Survey holds most of the references listed below, and copies may be obtained via the library service subject to copyright legislation (contact libuser@bgs.ac.uk for details). The library catalogue is available at: <http://geolib.bgs.ac.uk>.

BOWIE, R.C.; GILLANDERS, R.J. AND MORLEY, C. 2012. Single Onshore Borehole Index (SOBI) User Guide and Handbook – “The SOBI Bible”. *British Geological Survey*, pp112. (OR/12/006)

CRUMMY, J. 2008 Marine survey Data Management Handbook: BGS specifications. *Nottingham, UK, British Geological Survey*, 108pp. (IR/08/024). GRAHAM, C.C. AND HENNI, P.H.O. CONTRIBUTORS / EDITORS

G WILDMAN, K WESTHEAD, L AULT, P HENNI AND E CAMPBELL, Defining BGS digital materials in accordance to NERC Data Policy, *BGS Internal Report IR/11/056*