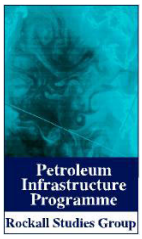


Project 98/28

Development of RSG Data Handbook and RSG Data Inventory

Final Report

Attachment 1: RSG Data Handbook

July 2000

John Gowen
Director
CSA Group

John Wallace
Director
Informatic Management

CSA Computing Services Ltd.
Marine Informatics Ltd.

TABLE OF CONTENTS

LIST OF APPENDICES

1 INTRODUCTION

The Petroleum Infrastructure Programme (PIP) was established by the Petroleum Affairs Division (PAD) of the Department of the Marine and Natural Resources in 1997. It was formed in conjunction with the award of exploration licences under the Rockall Trough Frontier Licensing Round.

The Rockall Studies Group (RSG) is a sub-group of the PIP and was set-up to address common industry problems in the Rockall Trough through:

- Regional data gathering Geology and Geophysics / Geotechnical / Environmental / Metocean
- Research projects, both applied and academic
- Scholarships associated with the research
- Research cruise sponsorship
- Provision of a forum to facilitate co-operation between industry and government

In 1999, the RSG members recognised the need for the introduction of guidelines on data handling, so that data consistency and effective data management could be achieved across the RSG programme. CSA Computing (CSAC) and Marine Informatics (MI) proposed a data management solution, with the design and implementation of a Data Handbook (RDH) and Data Inventory (RDI) for the RSG, to help it achieve some of its data management objectives.

Both CSA and MI have extensive experience in designing and implementing data management solutions to the natural resources sector and was the team that completed RSG Project 98/18 entitled *Technical Review of RSG Database Requirements*, which acted as an assessment of the data requirements for the RSG.

2 PROJECT OBJECTIVES

The objectives of the project were to design and implement Phase 1 and 2 of the database solution recommended in RSG Project 98/18. These were to create the;

1. RSG Data Handbook (RDH)
2. RSG Data Inventory (RDI)

The RDH describes data management procedures for the RSG projects, including how to submit, access, retrieve and request data. It gives guidance on data formats and other data management practices and policies. Any deliverable (in particular data sets) should have accompanying documentation (meta-data) so that the end user can fully utilise the data and be aware of any constraints or limitations.

The RDI describes the data sets collected by the various RSG-funded research projects, and is, as such, an inventory of meta-data. It does not contain any data, but instead will be used to direct users to the many sources. The data required to populate the inventory will be supplied by RSG project partners. A meta-data form was designed to collate the

meta-data. The RDI was developed to run on the Internet, allowing easy access by RSG members. The interface was developed to be user-friendly and is based on a MS-Access database initially but will be portable in the future to any of the well-known client server Relational Data Base Management Systems (RDBMS), such as Oracle, SQL Server or Sybase.

3 PROJECT BREAKDOWN

The activities involved in creating the RDH and RDI were viewed as continuous however CSA/MI proposed to divide the project into four distinct phases. The project life cycle was eight months.

In addition, it was CSA/MI's opinion that the RDH and RDI created from this project should have a wider focus than just the RSG data. It was therefore proposed to produce a data management solution that will be applicable to other Petroleum Infrastructure Programmes that are similar in nature to the RSG.

3.1 PHASE 1 – PRESENTATION AND REFINEMENT OF THE IMPLEMENTATION STRATEGY FOR RDH AND RDI

Phase 1 began with the presentation and refinement of the implementation strategy for the RDH and RDI. This was designed;

- to allow the key RSG members to give input into the data management solution, prior to the commencement of Phase 2 of the project

This involved sending a press release (see Appendix 1) to each of the RSG Project Managers. This document outlined the project aims and objectives and notified project managers that they would be contacted for certain RSG funded research project information.

3.2 PHASE 2 – DESIGN OF DATA HANDBOOK AND DATA PROTOCOL

3.2.1 Data Management Protocol

The RSG Secretariat has issued approximately thirty contracts. These projects come under four technical headings; Sub-surface (SSTC), Met-Ocean (MTC), Environmental (ETC) and Sea Bed (STC). The majority of these projects have yet to be completed and data has yet to be delivered to the RSG. The contracts for each of these projects do not have a Data Management Protocol (DMP) attached. DMPs are a data handling policy document standard in many industries, particularly where data is being collected, processed or interpreted. It was planned to design a DMP to be presented to each project manager as an addendum to existing RSG contracts. However on reflection it was thought that the DMP should form part of the RDH as to keep documentation to a

minimum and so facilitate full cooperation from project managers. It was planned that each Project Manager would be sent the RDH when completed.

3.2.2 Design Data Handbook

The RSG Data Handbook (RDH) describes the procedures for reporting on data collected from RSG projects. It is designed so that all the project data submitted to the RSG are described and controlled to an agreed and common standard. The information provided in the handbook is designed to act as a guide to the reader, on how data is to be documented prior to data submission to the RSG secretariat, how to write and document data output from projects, and how to fill-in the Meta-data submission forms (see Section 3.3).

The data deliverables described in RSG Project 98/18 were consulted when designing the RDH. This was to ensure that the Data Handbook covers all aspects of the RSG projects. New projects were checked that were not in RSG Project 98/18.

Now complete, the RDH will the give the data originator a detailed outline on the accompanying information needed when submitting data to the RSG Secretariat. As well as giving guidelines to the submission of experimental data, it will also give detailed description on submitting;

- Data Collection Reports
- Cruise Reports
- Sample/Data Processing Reports

Draft copies of the RDH have been sent to RSG Project Managers. The RSG Secretariat will distribute the final RDH document that accompanies this report.

3.3 PHASE 3 – DESIGN OF DATA DELIVERABLE MAP AND META-DATA FORM

3.3.1 Map of data deliverables on existing and upcoming projects

It was essential to have accurate information about the data being collected by each RSG project when designing the RDI. This information included a description of deliverables and their associated physical format. The RSG Secretariat was consulted to collate the information, and RSG project managers where appropriate. This work resulted in the production of a comprehensive map of all the data deliverables to the RSG. The map is called the Data Deliverable Map (DDM) and acted as the precursor to the RDI, see Appendix 2.

3.3.2 Design, distribution and completion of the Meta-data Form

Deliverables from the RSG Projects should have accompanying documentation (meta-data), which characterises and fully describes the data. Meta-data from each of the RSG projects will be used to populate the RDI. The Meta-data form captures the relevant meta-data to populate on RDI.

A Meta-data form was designed to capture the necessary meta-data and be simple for end-users to use and populate. Sample meta-data from ten key RSG projects were chosen to populate the RDI to show its functionality. Project managers were sent the Meta-data form with clear instructions as to how to fill it in. All ten project managers responded and the their associated meta-data currently forms part of the RDI. The meta-data received was of poor quality. Many of the projects had finished with deliverables submitted. There was no obligation on project managers to fill out the form comprehensively.

It was beyond the terms of reference of this project to populate the entire RDI. Meta-data received after the hand over date for this project will be managed by the RSG.

3.4 PHASE 4 – DESIGN OF RDI DATA BASE AND WEB-SITE

CSA/MI proposed a phased approach for the development of the RDI system. These can be summarised as follows:

- Project Planning and Specification
- Database Backend Design and Development
- Web Interface Design and Development
- Testing
- Implementation and Handover

A detailed technical description of the RDI database is presented in the *Project Proposal for Development of the RSG Data Handbook and RSG Data Inventory*, see Appendix 3. The RDI is now currently hosted for three months on the MI web server. The URL for the RDI is at http://www.informatic.ie/piprsg/ (Appendix 4).

4 SUMMARY

The RSG realised the need for effective data management with RSG Project 98/18 *Technical Review of RSG Database Requirements.* The diversity and number of projects led to the proposal of a phased data management solution:

> "The first stage is a Data Handbook (RDH) which is strongly recommended to provide guidelines for deliverables and sharing of data. The second stage is a web based index system (RDI), which will allow users to be directed to various data sources."

This two-phased approach has been adopted in this current project with the creation of the RDH and RDI. Both data management projects (98/18 and 98/28) have been completed after RSG contracts have been issued. Data has already been delivered to the RSG Secretariat. In this respect both data management projects can be seen to be retroactive. The meta-data form was sent to Project Managers of projects that had finished and there was no motivation for these people to populate the forms other than goodwill. Data guidelines were not given for many RSG projects and this resulted in many datasets being delivered in formats that may not facilitate use by interested parties. In hindsight it would have been more efficient if a data policy had been conceived with the inception of the RSG. Data management could have been in integral part of RSG contracts where final payment would have relied on certain data management criteria being realised. This is a useful and productive way to ensure data management compliance.

The RDI is an inventory of meta-data that is supplied by project partners. The RDI is only of use if it contains up to date, comprehensive meta-data. Many of the RSG projects are finished and there is no obligation on the partners to submit the necessary meta-data. The effectiveness of the RDI as meta-data resource is dependant on the quality of the meta-data that it contains. It was beyond the scope of this project to populate the RDI. It is therefore recommended that the RSG commission a contract for the population of the RDI.

The RDH and RDI created from this project have a broader application that simply for the RSG. This data management solution will be portable to similar Petroleum Infrastructure Programmes. The knowledge and experience of the both projects, 98/18 and 98/28, should be implemented with any new programmes where data will be collected and delivered. Data management should be a priority and project partners should be obliged to cooperate with any data management policy. This will ensure that data can be of use to interested parties and will be of benefit for future use.

98/28:Development of RSG Data Handbook and RSG Data Inventory

1.1 Introduction

The Rockall Studies Group (RSG) has been created under the Petroleum Infrastructure Programme (PIP), whose priority is to promote hydrocarbon exploration activities in Ireland.

RSG Project 98/18 was undertaken in 1998 as a joint partnership between CSA Computing Services Ltd. (CSA) and Marine Informatics Ltd. (MI) to assess the data requirements for the RSG. This project made several recommendations for a database solution, and as such, forms the basis for project RSG 98/28 where CSA and MI will design and implement Phase 1 and 2 of the database solution.

1.2 Project Methodology

The objectives of RSG Project 98/28 are to create:

- ***The RSG Data Handbook (RDH)***

The RSG Data Handbook will describe data management procedures for all the RSG funded research projects, including how to submit, access, retrieve and request data. It will give guidance on data formats and other data management practices and policies. The handbook will not necessarily supercede already existing procedures, but will ensure that all project data are described and documented to an agreed standard.

- ***The RSG Data Inventory (RDI)***

The objective of the RDI is to describe data sets (meta-data) collected by all the RSG funded research projects. It is intended that this will not contain any data but instead will be used to direct users to the many data sources. The meta-vdata required to populate the inventory will be supplied by RSG projects. It is envisaged that this user-friendly system will be hosted on the Internet, allowing updating and remote access.

1.3 Project Structure

CSA/MI propose to structure the project into four distinct phases;

- Presentation and refinement of the implementation strategy for RDH and RDI
- Design of a data handbook and data protocol
- Design of a data-deliverable map and meta-data form
- Design of the RDI database and website

1.4 Deliverables

There are four distinct deliverables;

- Final Report
- Data Handbook
- Meta-data form
- RDI Database/Website

The project team consists of Charlotte O'Kelly (MI), Niall Fahy (CSA), Mark Finucane (MI) and John Gowen (CSA). During the course of RSG Project 98/28 the project team will be contacting project managers and key personnel for certain RSG funded research project information.

THE RSG PROJECTS

RSG ADMINISTRATION/DATA ASSIMILATION

DATA REPROCESSING

NEW DATA ACQUISITION

A analogue data
D digital data
R raw data
P processed data
I digital images
M digital maps

98/19 Seabed Sampling using ships of opportunity
A

98/11 TTR7 data licenced to RSG
D(RPI)

97/14 TOBI Multiband imagery
A D(RPI)

97/14a TOBI processing and interpretation
A D(RPI)

97/11b RAPIDS 3 Quality control
A D(PI)

97/11 RAPIDS 3 Deep seismic
A D(RP)

97/11a Vessel HSE audits
A

98/1 Gravity and Magnetic Studies
A D(RPMI)

99/2 Multi Sensor Core Logging (MSCL) and core imaging
A D(I)

97/50a Biostratigraphy of Atlantic marine cores
A D

97/8 Fluid inclusion studies
A D

98/20 Geotechnical sample analysis
A D

98/21 Geochemical sample analysis
A D

98/6 Cetacean and Seabird Monitoring
A

97/50 BGS Atlantic Margin Drilling
A D(RPI)

99/1 Sediment budget and mass wasting on slopes flanking the Rockall Basin-slope architecture and internal stratigraphy
U

98/28 RSG data handbook and RSG data inventory
A D

98/18 Technical review of proposed database application
A D

98/22 RSG website development and maintenance
D

97/50b Provision of secure weblink and transmission link Chalenger and Norskaid
D

98/26 Environmental analysis of seabed samples from Rockall Trough
A D

97/34 High resolution biostratigraphy at the margin of the Rockall trough
A D

97/28 Sedimentological analysis of deep borehole cores
A D(RPI)

98/15 West of Ireland descriptive regional climatology
A D

97/29 MetOcean Strategy
A D(R)

98/23 Geohazard and shallow stratigraphy study
D(PMI)

97/3 Structural Nomenclature of elements in Rockall/ Porcupine area
A D(M)

97/40 Gravity and Magnetic Studies
A D

99/3 New methods for assessing thermal maturity (using samples from onshore and offshore samples)
U

97/51 Seabed data gathering and preliminary assessment
A D

97/52 Environment data gathering and preliminary assessment
A D

97/2 Statistical Characterisation of reflectivity patterns in deep seismic profiles
A D

97/21 Imaging through high impedence heterogenous layers
A D

97/16 Crystalline basement study
A D

DETAILED LIST OF DELIVERABLES FROM APPROVED RSG PROJECTS

Project	Deliverable Description	Formats
97/11 RAPIDS 3	Raw Seismic Data Processed Seismic Cruise Report Final Report	**Optical disk** **Q files from Seismic Handler on CD ROM (SEG Y available on request)** Hard copy/**Word** Hard copy/**Excel & Word**
97/14 TOBI acquisition	Raw navigation Raw imagery Mosaiced Image Cruise Report Cruise Logbook	**Navigation Text file** **Optical Disk** Hard copy/**PDF format** Hard copy Hard copy
97/50 Atlantic margin drilling	Digital Airgun and Sparker data Seismic tape logs Analogue seismic data Raw navigation data and log sheets Cleaned navigation data Gravity Core locations Qubit navigation print-out Echo-sounder print-out Oceans system Sound Velocity Probe data Certificate of calibration for SVP Ship board laboratory log book, Track chart for each site Stratigraphic summary Gravity Core Drill Core Core interpretations Photographic inventory of Logs	**Coda format on Exabyte tapes** Prints and **PC disk** Hard-copy Hard-copy **P1/90 format** **Excel Spreadsheet** Hard-copy Hard-copy **Floppy disk** Hard-copy Hard-copy Hard-copy Hard-copy Drill core Drill Core **Excel** **CD Rom**
98/6 Cetacean and sea bird monitoring	Seabird and Cetacean Report Cruise Report	Hard-copy Hard-copy
98/19 Sea bed sampling using ships of opportunity	Cruise Report Core Descriptions	Gravity cores Box cores
97/8 Fluid Inclusion studies of deep borehole cores	Fluid inclusion petrography final report	Hard-copy/**Excel & Word**
97/34 High resolution biostratigraphy	Biostratigraphic charts, paleoenvironmental charts, Interpretation report	Large paper charts compiled from StartaBugs software Hard-copy/**Excel & Word**
97/28	Core summary sheets,	Hard-copy

Project	Deliverable Description	Formats
Sedimentological analysis of deep borehole cores	Core sedimentology logs, Clay mineralogy, Photographic inventory of Logs	Hard-copy/**Illustator** **Excel spreadsheet** **CD Rom (same as 97/500**
97/50a Biostratigraphy of Atlantic marine cores	Onboard biostratigraphy report	**Specialist software (StrataBugs)** **Data files in ascii format**
97/50b Secure website and transmission link between Chalenger and Norskaid.	Design/maintenance of secure website Website hosting.	**HTML**
98/20 Geotechnical sample analysis	Geotechnical test results, Geotechnical characteristic report	Hard-copy report Hard-copy report/**Excel & Word**
98/21 Geochemical sample analysis	Interpretation report Data Charts	Hard-copy report/**Excel & Word** Hard-copy report/**Excel & Word**
97/14a TOBI – processing & interpretation	Clean navigation, Processed imagery, AO Maps, Final report	**Text file on CD-Rom** **Erdas Imagine format** Same as 97/14 Hard copy
97/3 Structural elements nomenclature	Nomenclature map, geological sections and report	Map/**EPS format** Map/**EPS format** Hard-copy report /**PDF format**
97/29 A MetOcean strategy for the Rockall Area	MetOcean data compilation report	Hard-copy report **Access database on floppy disk**
97/51 Seabed data gathering and preliminary assessment	Seabed data compilation report	Hard-copy report/**Excel & Word**
97/52 Environmental data gathering and preliminary assessment	Environmental data compilation report	Hard-copy report/**Excel & Word**
98/11 TTR7 Cruise license	Photos, charts, core logs and geochemical data of samples	**CD-Rom with Acrobat multi-media viewer**
97/16 Crystalline basement study	PhD Thesis	Hard-copy maps and report/**Excel & Word**
98/1 Gravity & Magnetic studies and 2D / 3D interpretation	Digital Atlas of 23 conventional and shaded relief maps Set of digital grids 3D whole crust model Set of digital grids of 3D model surfaces Final and interim reports	Hard-copy and **CGM format on CD-Rom** **Zycor, ascii or other format on CD-Rom** Hard-copy and **CGM format on CD-Rom** **Zycor, ascii or other format on CD-Rom** Hard-copy/**Excel & Word**

Project	Deliverable Description	Formats
97/2 Statistical characteristics of reflectivity patterns in deep seismic profiles	Basement Geological Maps and sections Software designed PhD Thesis	Unknown format Exabyte tape with software on it Hard-copy/**Excel & Word**
97/21 Imaging through high impedance heterogeneous layers	Final Report PhD Thesis	Hard-copy
97/40 Gravity & Magnetic studies	Maps & Data files PhD Thesis	Unknown Hard-copy
97/11a Vessel HSE Audits	To provide vessel inspection and safety audit for the RV Akademik Boris Petrov at Lubeck in connection with RSG Project 97/11 RAPIDS 3. On completion of the visits the outcome shall be discussed with C.Keiller of Enterprise Oil prior to finalising report. To provide vessel inspection, safety audit and HSE plan for the Akademik Boris Petrov at Kaliningrad and the Celtic Voyager in connection with RSG Project 97/11 RAPIDS 3. To meet the management team for the RADIDS 3 project and identify areas of concern in the safety management systems. When new systems are called for these should be developed in close collaboration with GeoPRO, UCD, DIAS and the Marine Institute. On completion of the visits the outcome shall be discussed with C. Keiller of Enterprise Oil prior to finalizing reports.	
97/11b RAPIDS 3 (Deep seismic) Boris Petrov charter, project management, cruise report	Processing and initial report. Preliminary interpretation report Interpretation report with seismic models and profiles Advanced interpretation report Final report Seismic sections and travel-time curves	Hard-copy report Hard-copy report Hard-copy report/**PDF format on CD Rom** Hard-copy report Hard-copy report Hard-copy report
98/15 West of Ireland Descriptive Regional Climatology	Final Report	Hard-copy report/**Excel & Word**
98/18 Technical Review of Proposed Database Application	Data Dictionary questionnaires Meta-data dictionary User Requirement questionnaires User Requirement Document Presentation of Project Results	Hard-copy report Hard-copy report Hard-copy report Hard-copy report Hard-copy report
98/22 RSG Website development and maintenance	Design/Maintenance of RSG Website Website hosting	**HTML**
98/23 Geohazard and Shallow Stratigraphy Study	Seismic Interpretation Charts, model files Charts, sheet files	**Seisworks (Exabyte)** **DGN & CGM (CD ROM)** **DGN & CGM (CD ROM)**

Project	Deliverable Description	Formats
	Charts, plot files Digital terrain model Cross sections Data examples	**HP/GL2 or PetraSys (CD ROM)** **Space delimited ASCII xyz (CD ROM)** **DGN & CGM/GIF or JPEG (CD ROM)** **GIF or JPEG (CD ROM)**
98/26 Environmental Analysis of Seabed samples from RT	Interim and Final Report	Hard-copy report/**Excel & Word**
98/28 RSG Data Handbook and RSG Data Inventory	Final Report Data Handbook Meta-data form RDI-Functional Specification Document RDI Backend databases User Guide and Technical Documentation **RDI Website**	Hard-copy report/**Excel & Word** Hard-copy report/**PDF format on the RSG Website** Hard-copy report/**Executable (exe) format on the RSG Website.** Hard-copy report/**Excel & Word** **Access Databases** Hard-copy report/**Excel & Word**
99/1 Sediment budget & mass wasting on slopes flanking the Rockall Basin (slope architecture and internal stratigraphy)	PROPOSED (NO CONTRACT AS YET)	
99/2 Multi-Sensor Core Logging (MSCL) & core imaging	Multi-Sensor Core Logging (MSGL) and Imagery data Interpretation of MSCL data	**Illustrator format on CD Rom** Hard copy report/ **Excel & Word**
99/3 (NO DETAILS TO DATE) New methods for assessing thermal maturity (using samples from onshore and offshore Ireland)	PhD Thesis	

Technical Specification Document

Version: 1.1

Brian McNamara, Informatic Software Limited

REVISION CHART			
Version	**Primary Author(s)**	**Description of Version**	**Date Completed**
1.0	Brian McNamara	Initial draft	28th January 2000
1.1	Brian McNamara	Removed Web URL Links and Feedback	1st February 2000

Contents

1. Introduction

1.1 Background

This scope of this project is to develop a web based RSG Data Inventory, under the PIP programme.

1.2 Objective and Scope

The objective of this document is to provide interface mockups of what the application will appear like to the end user, and to describe the technical architecture of the application

1.3 Circulation List

- CSA
- IMI/Informatic Software Limited

2. Technical Architecture

2.1 Description of the Environment

The application will be developed using the following technologies:

- Server Side Technology
  - Intel PC machine (300MHz minimum), 128MB RAM (minimum), 2GB HDD
  - Cleaned Down Machine (i.e. fresh installation of NT Server 4.0, and then IIS 4.0) - in particular, no previous installation of ADO libraries
  - IP address accessible outside the firewall/proxy server (e.g. 194.125.54.xx)
  - Windows NT Server ver 4.0 running **only** IIS 4.0 on the default http TCP/IP Port (i.e. not a shared machine with O'Reilly Website)
  - Microsoft Active Server Pages (in VBScript)
  - ActiveX Data Objects
  - ODBC
  - Microsoft Access Configuration Utility (MDE file)

- Client Side Technology
  - Microsoft Internet Explorer Version 4.0 browser with Cookies enabled
  - Microsoft Outlook Express Newsgroup Reader
  - 1024 x 768 Screen Resolution
  - JavaScript
  - VBScript
  - Java

3. Programs and Procedures

3.1 List of Programs/Procedures

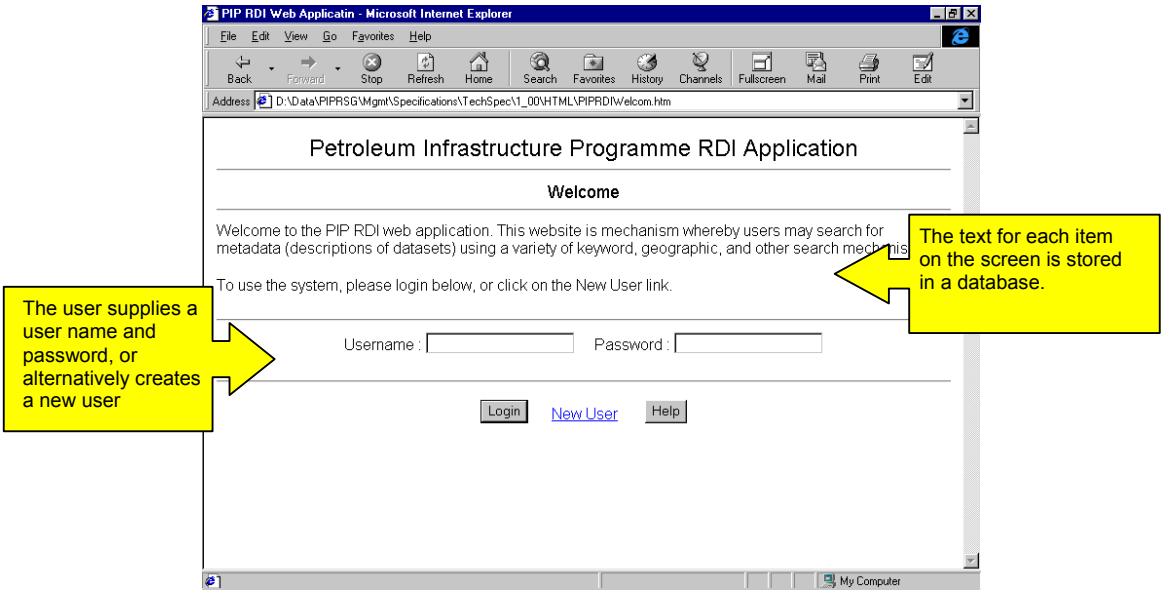
The following procedures form the modules of the RDI Web application.

Module	Technology	Description
ISLMultiLingualMeta Database	MS Access relational database	Provides the content for the database driven application. Also includes definitions for interface components.
ISLMultiLingualMeta Manager	Server side VB Script	The ASP pages that dynamically build the interface on-the-fly from the relational database
ISLMetaConfigTool	MS Access MDE application	Configures the application (e.g. keywords, administrative monitoring of added data by users)
ISLFormVal	Client Side JavaScript	Form validation
ISLMapIt	Client Side Java	Mapping APPLET

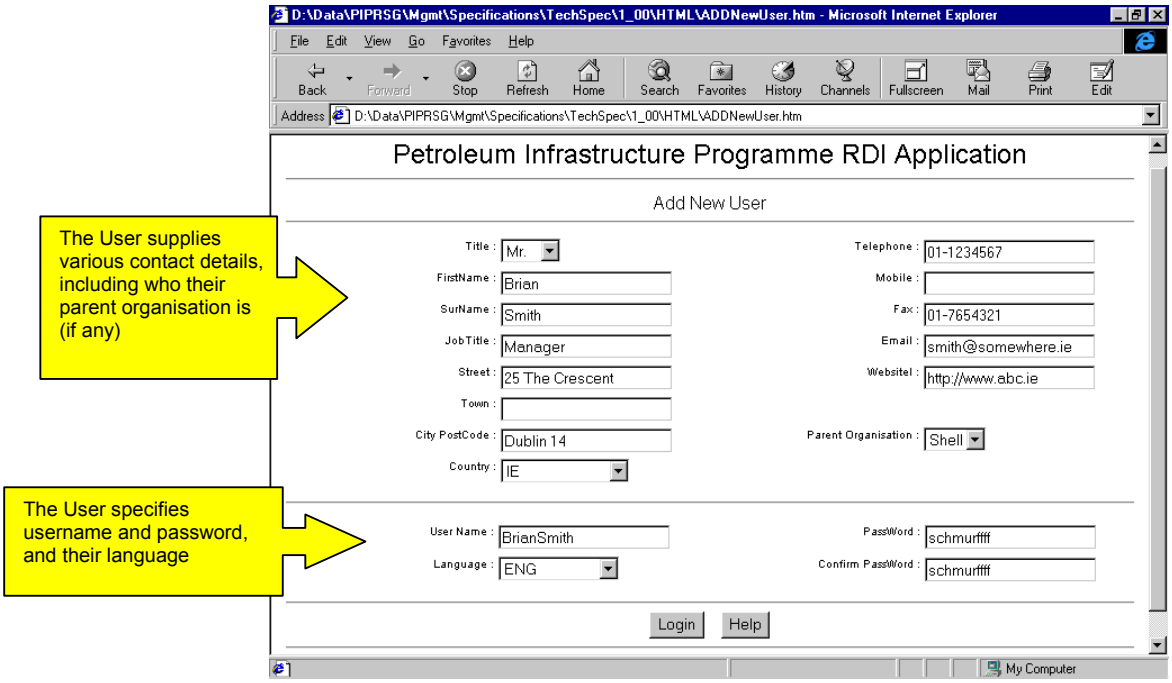
4. Screen Layouts

4.1 Initial Security Screen

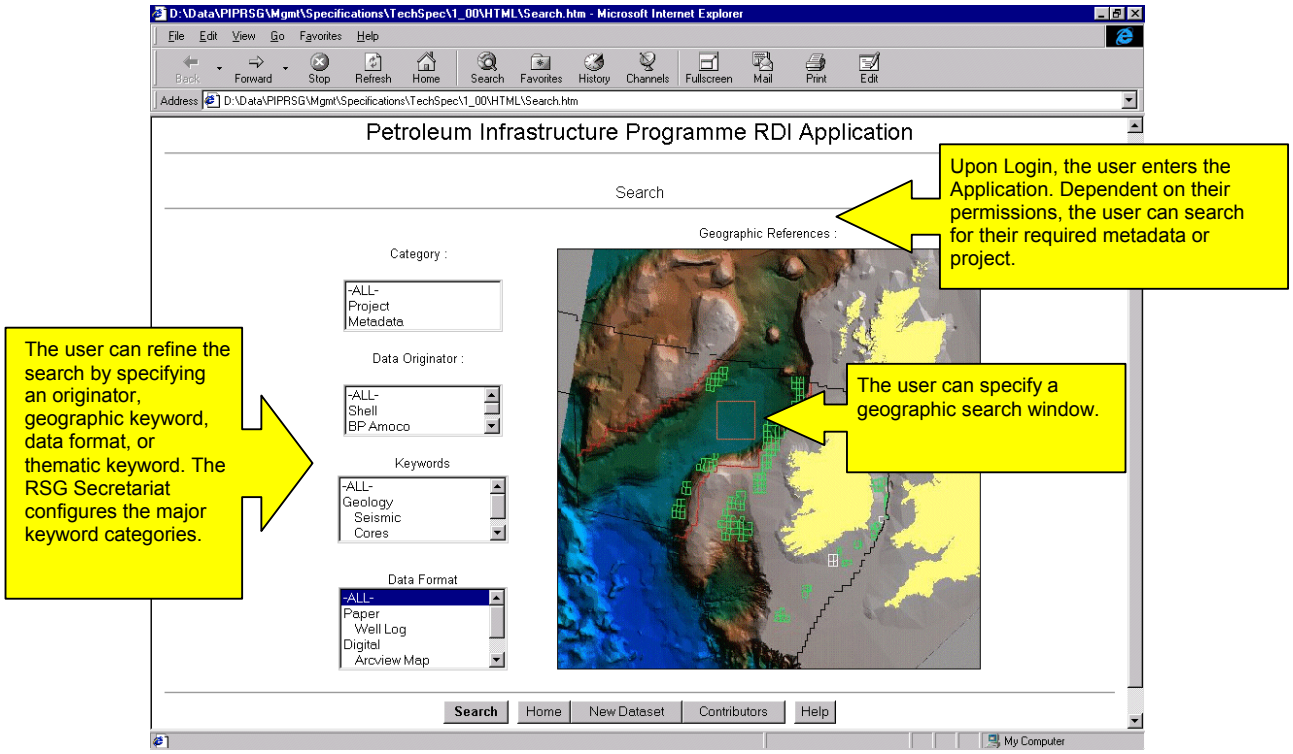
4.2 Title Page and Login Screen



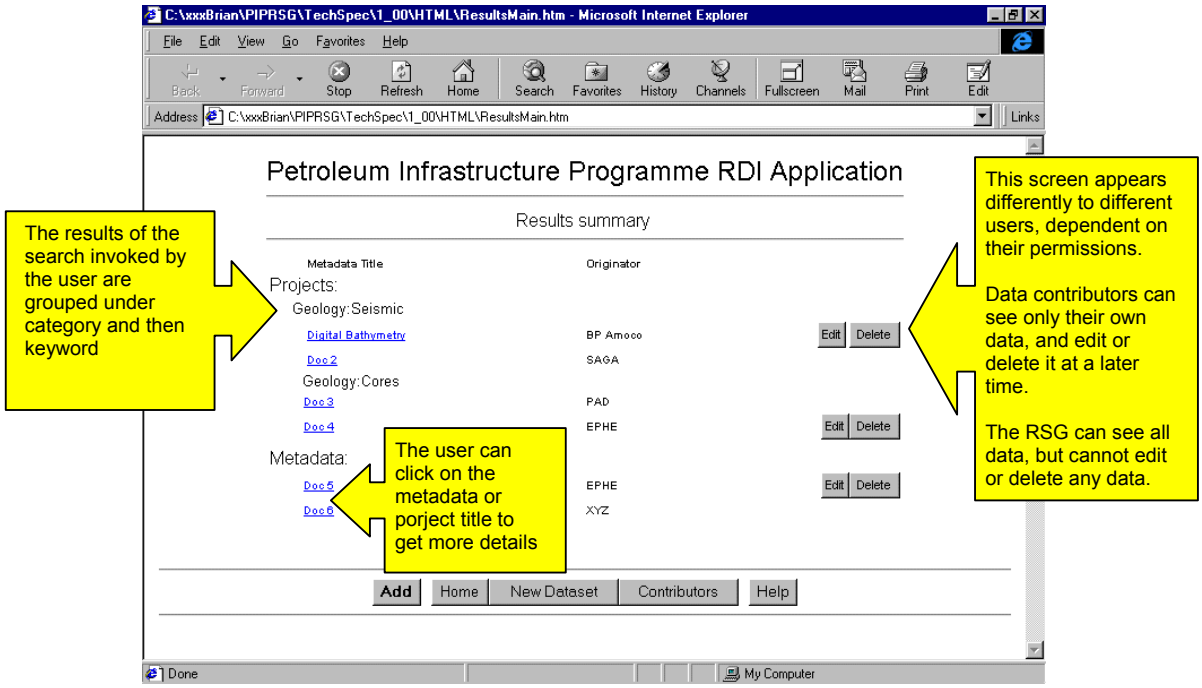
4.3 Add New User Screen



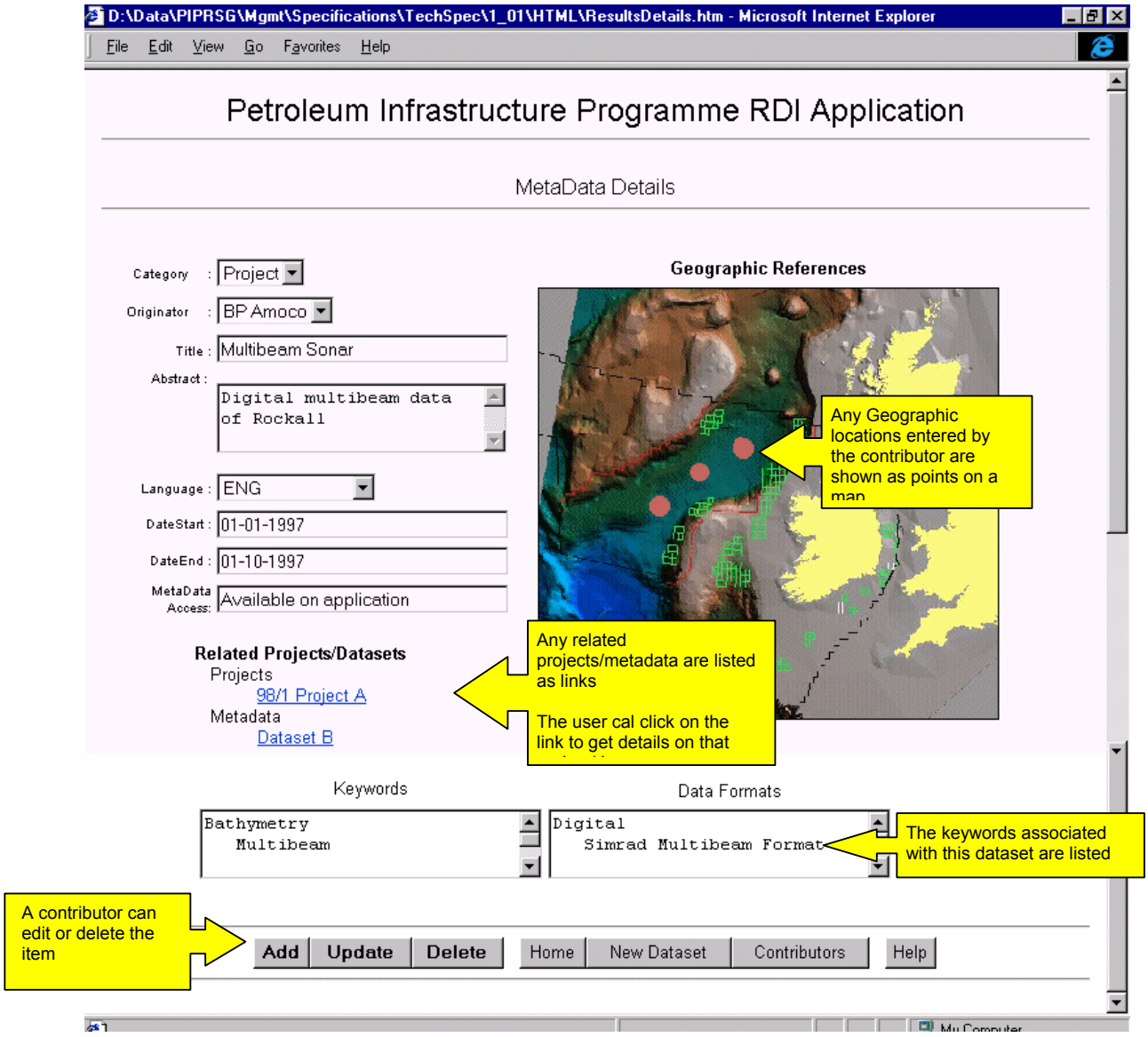
4.4 Search Screen (Home Screen)



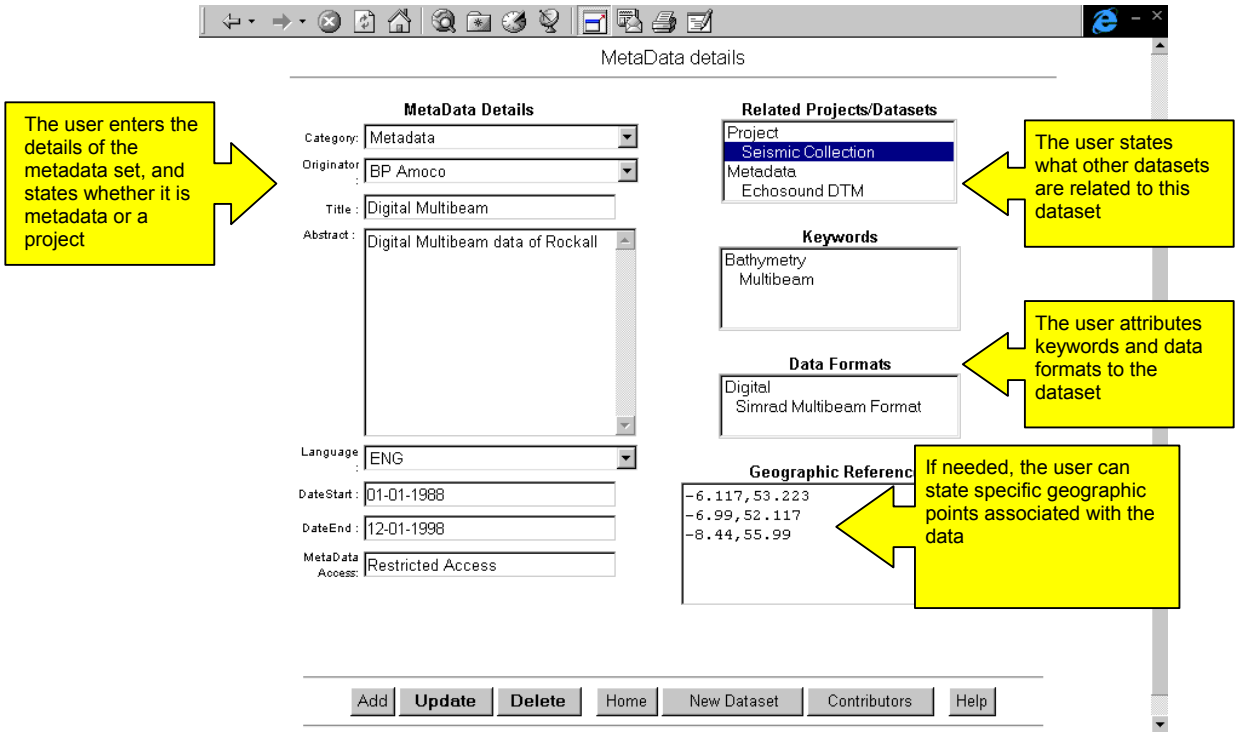
4.5 Search Results Summary



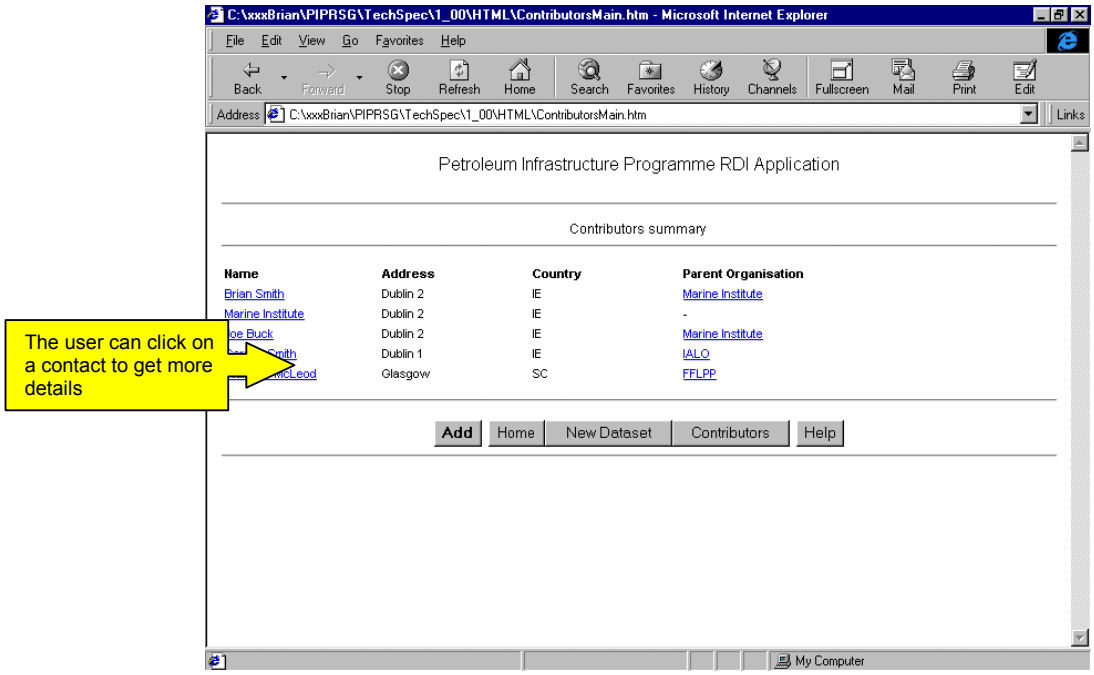
4.6 Metadata Set Details (Read Only)



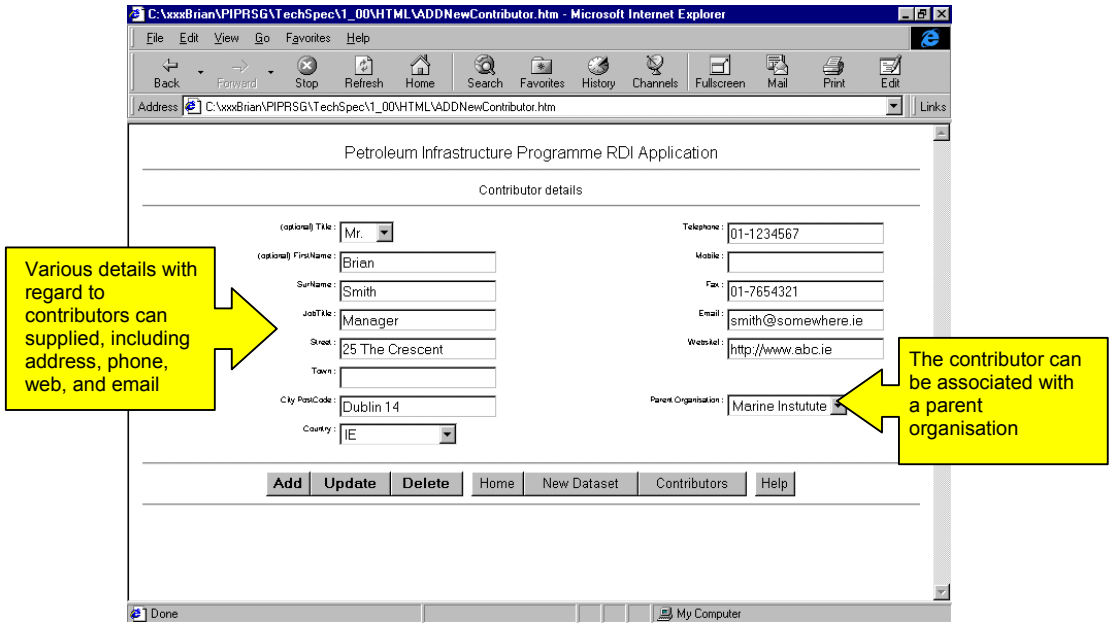
4.7 Metadata Entry Screen



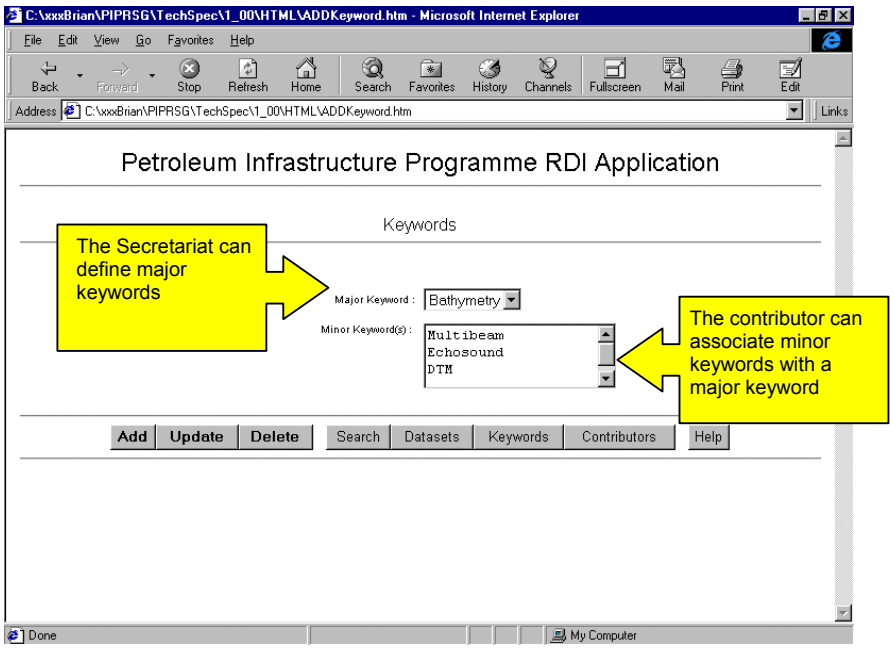
4.8 Contacts Summary



4.9 Contact Details



4.10 Keyword Definitions



4.11 Help Page

An application help page shall also be provided which shall include concise instructions on how to use the system.

5. Operational Procedures

5.1 Backup Procedures

The Client should ensure that the system in included in its backup rota, and that both onsite and offsite copies are maintained.

6. Security

6.1 System LogIn

The system requires that any users of the system register their contact details, username and password upon first using the system.

The overall application is also protected by a global password set in IIS.