

MARINE GEOLOGY DATA ACCESSIBILITY IN THE EUROPEAN FRAMEWORK: THE I.G.M.E. PARTICIPATION IN THE GEO-SEAS PROJECT

Zananiri I.¹, Mitropoulos D.¹, Zimianitis V.¹, Ioakim Chr.¹, Papadopoulos V.¹ and Efthimiou G.¹

¹*Institute of Geology & Mineral Exploration, National Centre for Sustainable Development, 1 Spirou Loui str., 13677 Acharnae, Greece, izanan@igme.gr, mitropoulos45@gmail.com, zimianitis@igme.gr, ioakim@igme.gr, rootmail@igme.gr, efthimiou@igme.gr*

Abstract

During more than 30 years of activity and participation in numerous research projects, the Institute of Geology and Mineral Exploration of Greece has carried out extensive work in marine geology and a vast amount of data has been collected. In order for those data to be available and used by the international scientific community, a need for standardization and harmonization was imminent. Towards this scope IGME, along with 25 other marine geological and geophysical data centres, located in 17 European maritime countries, participated in the GEO-SEAS project: the implementation of an e-infrastructure. The aim of the project was to enable users to identify, locate and access pan-European, harmonised and federated marine geological and geophysical datasets and derived data products held by the data centres through a single common data portal. All available data have been harmonized according to the GEO-SEAS guidelines and standards. Detailed metadata files were created, accompanied by the corresponding data and additional information files. The standardization, harmonization and provision of data and metadata, carried out through the GEO-SEAS project, are aligned with European directives and recent large-scale framework programs on global and European scales.

Key words: *e-infrastructure, standardization, harmonization, metadata.*

Περίληψη

Το Ινστιτούτο Γεωλογικών και Μεταλλευτικών Ερευνών, κατά τη διάρκεια περισσότερων από 30 χρόνων συμμετοχής σε διάφορα έργα και ερευνητικά προγράμματα, έχει σημαντική δράση στον τομέα της Υποθαλάσσιας Γεωλογίας, με αποτέλεσμα τη συλλογή ενός τεράστιου αριθμού δεδομένων. Για να καταστούν τα δεδομένα αυτά διαθέσιμα στην διεθνή επιστημονική κοινότητα ήταν επιτακτική η ανάγκη τυποποίησης και εναρμόνισής τους. Για το σκοπό αυτό το ΙΓΜΕ, μαζί με 25 ακόμη κέντρα, από 17 χώρες, που διαθέτουν γεωλογικά και γεωφυσικά δεδομένα από τον θαλάσσιο χώρο, συμμετείχε στο έργο GEO-SEAS που αφορούσε στη δημιουργία μιας ηλεκτρονικής υποδομής. Οι δράσεις του έργου στόχευαν στο να έχουν οι τελικοί χρήστες τη δυνατότητα αναζήτησης και πρόσβασης σε πανευρωπαϊκά, εναρμονισμένα και σύμφωνα με τα διεθνή πρότυπα δεδομένα υποθαλάσσιας γεωλογίας, μέσω μίας ενιαίας διαδικτυακής πύλης. Όλα τα διαθέσιμα δεδομένα του ΙΓΜΕ εναρμονίστηκαν με βάση τους κανόνες και τα πρότυπα που ορίστηκαν στα πλαίσια του έργου GEO-SEAS. Παράχθηκαν λεπτομερή

*αρχεία μεταδεδομένων που συνοδεύονται από τα αντίστοιχα αρχεία δεδομένων. Η τυποποίηση, εναρμόνιση και διάθεση των δεδομένων, όπως καθορίστηκε μέσα από το έργο GEO-SEAS, είναι σε πλήρη συμφωνία με τις Οδηγίες της Ευρωπαϊκής Ένωσης καθώς και τα πρότυπα που χρησιμοποιούνται σε άλλα προγράμματα διάθεσης δεδομένων σε Ευρωπαϊκές και διεθνείς κλίμακες.
Λέξεις κλειδιά: ηλεκτρονική υποδομή, τυποποίηση, εναρμόνιση, μεταδεδομένα.*

1. Introduction

Marine research in the previous decades has been extensive throughout Europe and has led to the collection of an enormous amount of data. However, the marine observing system is highly fragmented: in the countries bordering the European seas more than 600 scientific data collecting laboratories from governmental organizations and private industry have been identified. They collect data by using various sensors on board of research vessels, submarines, fixed and drifting platforms, airplanes and satellites, to measure physical, geophysical, geological, biological and chemical parameters, biological species etc. The collected data are neither easily accessible, nor standardized. They are not always validated and their security and availability have to be insured in the future.

In Greece, a country of 15,147 km coastline and a large surface of territorial sea, numerous studies on marine geology and related disciplines have been carried out over the past decades. These studies, implemented in the frame of national and international research projects by the Institute of Geology and Mineral Exploration, the National Centre of Marine Research, the Laboratory of Marine Geology and Physical Oceanography of the University of Patras, the Department of Marine Sciences of the University of Aegean and other collaborating institutes, have resulted in the production of a large set of diverse marine data.

Access to marine data is of vital importance for marine research and a key issue for various studies, from the climate change prediction to off shore engineering. Thus, the need for standardization, harmonization and unhindered search – and subsequent conditional provision – was imminent. Towards this scope, the GEO-SEAS project, an Integrated Infrastructure Initiative (I3) of the Research Infrastructures programme within EU FP7 (contract number RI-238952), was implemented

The data populated through GEO-SEAS concern marine geological and geophysical data sets.

This paper aims at providing an overview concerning (1) the main achievements of the GEO-SEAS project, (2) the organization of IGME marine geology data and metadata for inclusion in the project Common Data Index.

2. The Geo-Seas Project: an Overview

The pan-European GEO-SEAS project (“Pan-European infrastructure for management of marine and ocean geological and geophysical data”, see also <http://www.geo-seas.eu>) was launched in May 2009, having the duration of 45 months, and ended in 31st January 2013. It was materialized by 29 partners – comprising 25 geological and geophysical data centres and 4 institutes that bring in additional expertise – from 17 European countries, bordering the North-East Atlantic and its adjacent seas: the Mediterranean, the Black Sea, the Baltic and the North Sea. The project coordinator was the British Geological Survey – National Environment Research Council (UK), while Marine Information Service MARIS BV (Netherlands) was appointed as the technical coordinator.

The aim of the Geo-Seas project was to enable users to identify, locate and access pan-European, harmonised and federated marine geological and geophysical datasets and derived data products held by the data centres through a single common data portal. Thus, the main objectives were:

- To expand the SeaDataNet (<http://www.seadatanet.org>) infrastructure by including the network of national geological surveys and selected geological research institutes.
- To publish and maintain a common catalogue and index of available data, data products and services, managed by national geological surveys and research institutes in Europe.
- To harmonise quality standards and exchange formats in order to facilitate access and delivery of these data to the various user communities, from research, government and industry.
- To provide federated, marine geological and geophysical datasets, ready for use in the user communities using middleware to access and deliver data via the internet to their networks and grid infrastructures.
- To develop and to provide new dedicated geological and geophysical data products and services.
- To formulate, determine and facilitate arrangements for use of these data and data products.
- To formulate, determine and facilitate arrangements for consolidation and long-term exploitation of the Geo-Seas infrastructure.
- To promote the infrastructure and its services to users and to encourage adoption of protocols, standards and tools by other marine geological and geophysical institutes and companies.

In order to fulfil these tasks, the project was organized in 11 work-packages involving management, training, e-infrastructure building, data and metadata records production, dissemination and software development activities.

3. The IGME Database

Marine Geology at IGME is currently carried out within the activities of the Division of General Geology and Geological Mapping. The marine geology team comprises a group of scientists (geologists, geophysicists and palaeontologists) and laboratory staff and is assisted by the IT team. The infrastructures comprise field equipment for bathymetry and seismic reflection measurements, sediment grabbers and cores and 2 laboratories, the Marine Geology Laboratory (grain-size analysis etc.) and the Stratigraphy-Palaeontology-Palynology Laboratory, complemented by other disciplines within IGME (e.g. Mineralogy, Geochemistry, GIS mapping etc.).

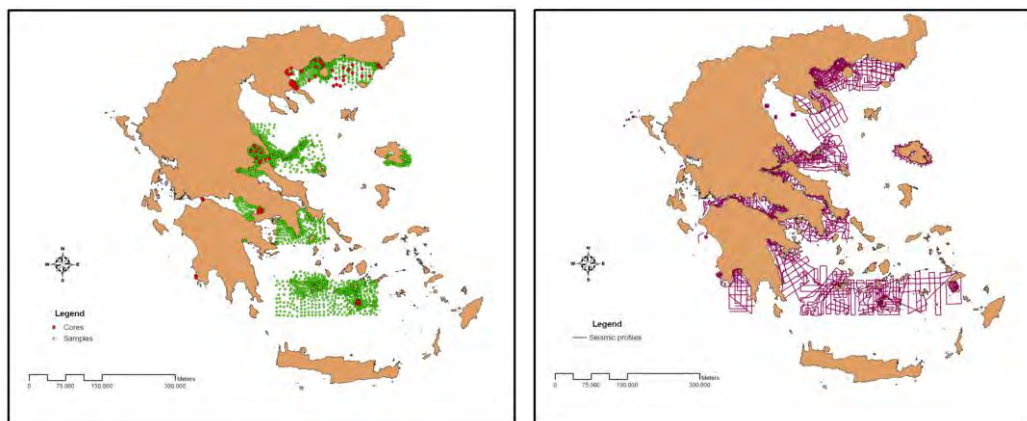


Figure 1 - The dataset of IGME included in the Geo-Seas portal: sediment grab samples and sediment cores (left), lines of seismic profiles (right). These data have been collected by IGME from 1979-present.

During more than 30 years of activity and participation in numerous research projects, the Institute of Geology and Mineral Exploration of Greece has carried out extensive work in marine geology. As a result a vast amount of data has been collected, including marine sediment samples and cores, shallow and medium penetration seismic profiling, bathymetry and side scan sonar data. For the scope of the GEO-SEAS project the IGME database was updated and completed with all the information required. Metadata and associated data files from a total of 2147 sediment grab samples, 79 sediment cores – comprising 437 specimens – and 547 seismic profiles (~20,000 km) – reaching a total of 2603 scanned images – were prepared for inclusion in the GEO-SEAS e-infrastructure (Figure 1).

3.1.1. Metadata Records

The Common Data Index (CDI) is the central discovery service which enables users to have a detailed insight of the availability and geographical extent of marine data archived at the connected data centres. One of the primary goals of the project was to produce the CDI metadata records for inclusion in the GEO-SEAS infrastructure. For their creation the SeaDataNet standard (ISO19115 compliant) for geological and geophysical data was adopted and adapted to suit the GEO-SEAS needs and Common Vocabularies were expanded.

Table 1 – Definition of metadata characteristics.

MANDATORY	OPTIONAL
Dataset-id	Time resolution
Dataset-name	Vertical element (min-max observation depth)
Coordinates (or bounding box)	Vertical datum
Object type	Water depth
Heading 3	Datum of coordinate system
Period of sampling	Track resolution
Parameters measured	Polygon (for curve/surface data)
Brief summary of the content	Sampling interval
Sampling instrument	Cruise information (name – id – date)
Sampling platform class	Station information (name – id – date)
Data holding centre	Project title
Originator centre	
Distributor	
Collate centre	
Distribution websites & services	
Dataset access restriction	
Data format	
CDI file revision date	

The CDI population and maintenance activities at IGME included various sub-activities (Figure 2):

- Analysis of the formats and local availability of metadata
- Editing of existing metadata according to the GEO-SEAS standards and guidelines

- Compiling and validating a first test-batch of metadata
- Compilation and/or updating of the metadata and validation into a test site by MARIS
- Importing geological and geophysical XML metadata entries into the central CDI directory.

The generation of the CDI records was accomplished by using the MIKADO software, a tool written in Java, which enables data centres to prepare XML metadata files. In total 2773 CDI records were prepared to describe the IGME geological and geophysical data, including all the mandatory and optional information –where applicable. Moreover, to provide additional information for the available seismic profile data, xml metadata files describing the sampling methodology and seismic section data record were created, again using the MIKADO tool:

- 2603 SensorML metadata records, one for each individual segment of the various seismic sections, with information about: applied seismic method, dimensionality, data product, quality control, seismic source, receiver, acquisition and bandwidth.
- 547 O&M (Observation and Measurements) metadata files, one for each seismic section, with reference to the section bounding coordinates and links to all the SensorML metadata files for this section.

Finally, to conform to the low-resolution viewing services pre-requisite, low-resolution thumbnails of the 2603 scanned seismic profile images were created.

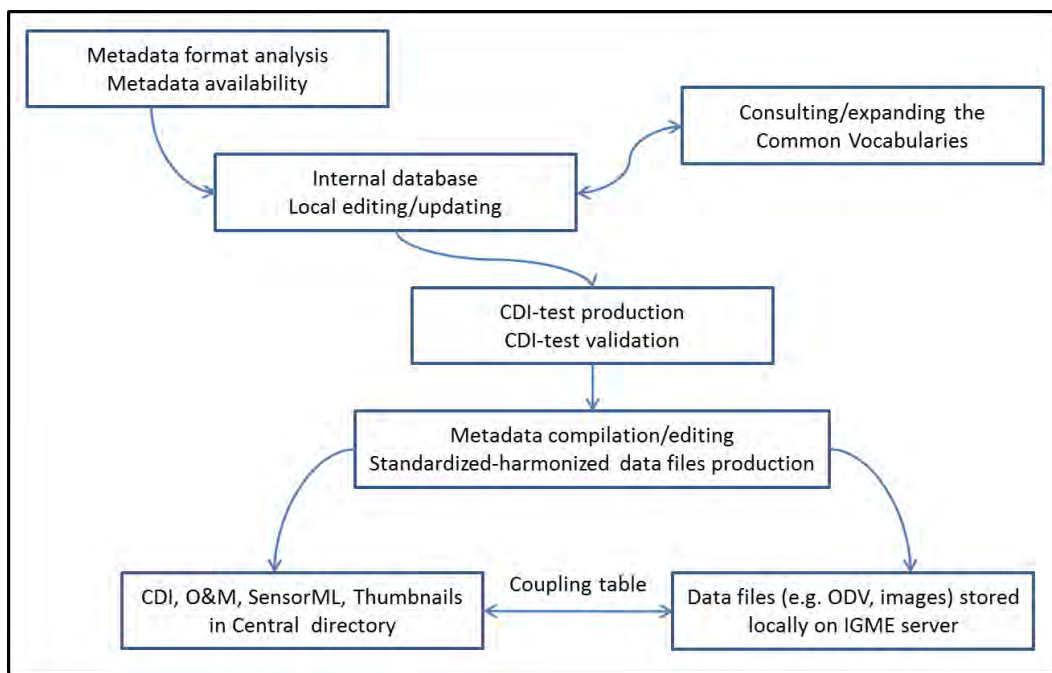


Figure 2 - CDI population activity at IGME.

3.1.2. Data Files

A necessary condition for the inclusion of metadata records in the CDI directory was the direct connection with the, locally stored, associated data records. It is well-known that geological data comprise a wide range of parameters, estimated visually and measured in the field or in the laboratory. Thus, to ensure that these records comprise a pan-European, easy to use, dataset, it was necessary to follow certain standardization and harmonization guidelines, which improved the consistency in the use of data in terms of meaning and format. Moreover, the data sets had to be delivered in the agreed exchange formats, common to all data centres.

Following the GEO-SEAS standardization and harmonization protocol, 2226 files – 2147 sediment grab samples and 79 sediment cores – of lithology and grain-size parameters were prepared. The files were in Ocean Data View (ODV) format, which is a commonly used INSPIRE-compliant transfer format. The generation of the ODV records was accomplished by using the NEMO software, a tool written in Java, which enables the conversion from any type of ASCII format to the ODV format. Each data parameter is defined using codified terms from Common Vocabularies, which were selected as the most appropriate and applicable to all participating data centres. For example, to describe main lithology many different classification systems have been used in the past are currently being used worldwide; in the scope of the GEO-SEAS project the Adjusted Folk Classification [Folk (1974) classification with additions by Flemming (2000)] was implemented. An example of how data input from an ASCII file is imported and defined using NEMO software is presented in Figure 3.

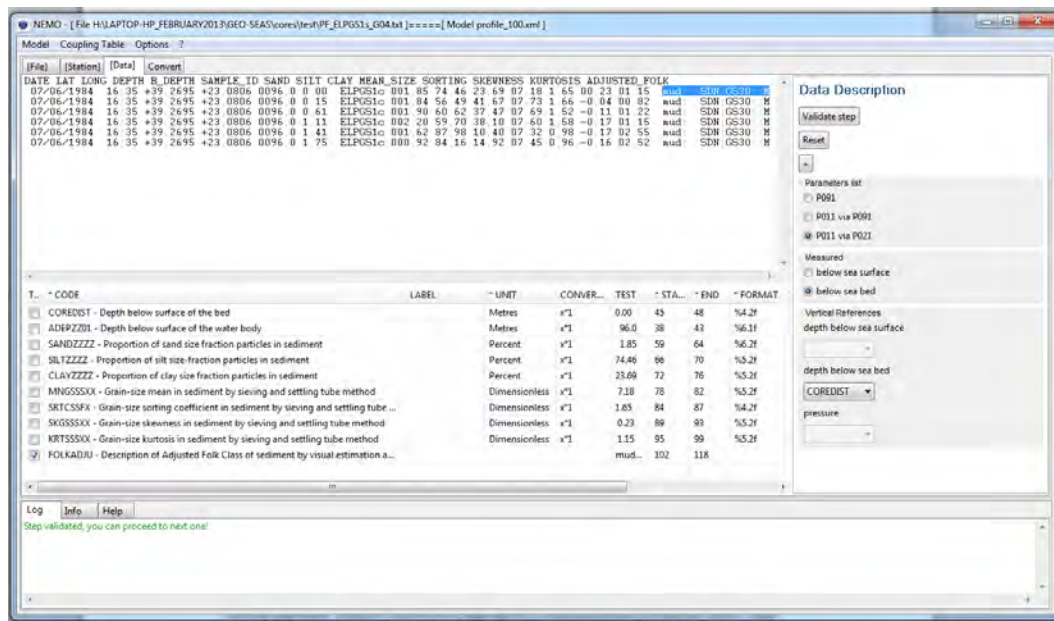


Figure 3 – Definition of data input for the preparation of ODV using NEMO software.

Concerning the geophysical data, the transfer format that was applicable to IGME was high-resolution image files with raw data representation. To achieve this all the seismic profiles, still in analogue format, were scanned to the required resolution and together with the ones previously scanned during SEISCANEX project (<http://www.noc.soton.ac.uk/gg/SEISCANEX>) they comprised a dataset of ~20,000 km of seismic lines, with a total of 547 seismic lines – amounting to a total of 2603 high-resolution images.

Finally, detailed records of palynology, palaeontological, radiochronological and other data on certain cores are being prepared in accordance with the GEO-SEAS policy and will be included in the e-infrastructure.

4. The Geo-Seas e-infrastructure

As previously mentioned the main purpose of the GEO-SEAS project was to provide a transparent method of data search and delivery. Towards this scope, the architecture and methodologies developed by the SeaDataNet project were adopted for the establishment of the GEO-SEAS infrastructure and portal. There are three basic elements in this architecture:

- An online discovery search engine (http://www.geo-seas.eu/v_cdi_v2/browse_step.asp) – the so-called Common Data Index (CDI) – with advanced search options based on the metadata provided, and an integrated data order service. Through this portal low-resolution viewing services are available for the seismic profiles.
- The actual delivery of data sets is done between the user and the selected data centre, using automatic approach – for data with no restrictions – or manual approval/denial in the case of restricted use/commercial data.
- The connection between the metadata records, uploaded in the CDI, and the actual data files is carried out with the use of the Download Manager, a Java component that handles all communication between the data centre system and the Request Status Manager service.

Through, this e-infrastructure, the Geo-Seas data, data products and services can be easily accessed and used by the following sectors: environmental research and monitoring; academic research; government; national and regional agencies; dredging; marine hydrocarbons; beach nourishment; land reclamation; sustainable energy; civil engineering (pipelines, offshore construction, aggregates); communications (submarine cables); shipping; fisheries; tourism; health. At the same time data centres still have secure data authorship and control over data provision and use.

5. Conclusions and Future Prospects

Through the implementation of the GEO-SEAS project the existing marine geological and geophysical information from IGME and 24 more data centres has been standardized, harmonized and is now available for search and download through a single portal. This has facilitated the search and use of the information from the scientific community and other users. Until now the recorded statistics are more than satisfactory and several data requests have been submitted from institutes, research centres, universities as well as private companies (e.g. for submarine cable suitability studies).

The standardization, harmonization and provision of data and metadata, carried out through the GEO-SEAS project, are aligned with European directives and recent large-scale framework programmes on global and European scales, such as GEOSS and GMES, EMODNET and INSPIRE. Thus, they allow IGME to fully conform to the EU directives for the provision of marine geology data.

The European Committee, through its Directives and general Policy on metadata and data accessibility promotes the pan-European approach on coherent management and data/data products provision. Thus, the maintenance of the GEO-SEAS portal, the inclusion of additional data, as well as the creation of pan-European marine geology data products, constitutes the target of future research in the e-infrastructures field.

6. Acknowledgments

We greatly acknowledge the European Commission for funding the GEO-SEAS project through the Integrated Infrastructure Initiative (I3) of the Research Infrastructures Programme (Research Project Number: 238952).

For the carrying out of the project we greatly acknowledge and thank all cooperating scientists, in Greek and European Institutes and Universities throughout the work in all these years. GEO-SEAS consortium partners: NERC-BGS (United Kingdom), NERC-BODC (United Kingdom), NERC-NOCS (United Kingdom), MARIS (Netherlands), IFREMER (France), BRGM (France), TNO (Netherlands), BSH (Germany), IGME (Spain), INETI (Portugal), IGME (Greece), GSI (Ireland), BGR (Germany), OGS (Italy), GEUS (Denmark), NGU (Norway), PGI (Poland), EGK (Estonia), LIGG (Lithuania), IO-BAS (Bulgaria), NOA (Greece), CIRIA (United Kingdom), MUMM (Bel-

gium), UB (Spain), UCC (Ireland), EU-Consult (Netherlands), CNRS (France), SHOM (France), CEFAS (United Kingdom), and LU (Latvia).

7. References

- Diviacco P. 2010. Report on standards to be used within Geo-Seas while transferring/sharing Geophysical data, *GEO-SEAS WP4 Report*, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, 1-37pp. Available online: http://www.geo-seas.eu/content/content.asp?menu=0330000_000000
- Flemming B.W. 2000. A revised textural classification of gravel-free muddy sediments on the basis of ternary diagrams, *Continental Shelf Research*, 20, 1125-1137.
- Folk, R.L. 1974. The petrology of sedimentary rocks, *Hemphill Publishing Co.*, Austin, Texas, 182 pp.
- MIKADO (Version 2.5) Software, 2013. Copyright IFREMER. Retrieved from: <http://www.seadatanet.org/Standards-Software/Software/MIKADO>
- NEMO (Version 1.4.5) Software, 2013. Copyright IFREMER. Retrieved from: <http://www.seadatanet.org/Standards-Software/Software/NEMO>
- Van Heteren, S. 2010. Standards for formats for transport of geological data, *GEO-SEAS WP4 Report*, TNO Geological Survey of the Netherlands, 1-28pp. Available online: http://www.geo-seas.eu/content/content.asp?menu=0330000_000000