International Conference on Marine Data and Information Systems (IMDIS 2016)
Gdansk (Poland) - October 11-13

Facilitating marine data discovery, representation and retrieval through metadata interoperability

Pablo Otero, Elena Tel and Gonzalo González-Nuevo*

Instituto Español de Oceanografía, Spain
***pablo.otero@md.ieo.es**

In Spain, the *Instituto Español de Oceanografía* (IEO; www.ieo.es) maintains the National Oceanographic Data Center (NODC), responsible for the compilation, storage and distribution of marine data. Thanks to the guidelines of standardization, the NODC serves geospatial information of the projects, vessels and observatories through large international frameworks like SeaDataNet and EMODNET. However, the navigation through these portals can become complex for the non-expert and non-English speaking end-users. To facilitate marine (geospatial and non-geospatial) data discovery to the Spanish community, a GeoNetwork catalogue has been configured. Through a XSL transformation, and following the guidelines of standardization, the catalogue also integrates Cruise Summary Reports (CSR) and Common Data Index (CDI) XML files created with MIKADO software.

INTRODUCTION

Access to marine data is of vital importance for traditional marine researchers, but also for a variety of professionals who use these data to tackle problems related to climate change, coastal engineering, fishing or aquaculture, among others. In addition, the demand of this kind of data by the general public is becoming more and more common (recreational navigation, nautical sports, tourism, etc.). Unfortunately, marine datasets are usually stored in specialized portals and they are often not indexed for Google, and will therefore not show up in end-user search results. Moreover, the navigation through these specialized portals is usually complex for the non-expert public.

Data producers interested in targeting a wider and general public, should design user friendly sites, in native language, with the use of the proper terminology and prioritizing end-user interests [1]. From an organizational point of view, to achieve this goal implies to walk a long path where standardization and interoperability are main steps for managing the large and diverse datasets collected by the oceanographic fleets and the automatic observation systems. To facilitate data discovery through a catalogue is a first step. Moreover, an additional effort must be done to accomplish the INSPIRE Directive (2007/2/EC).

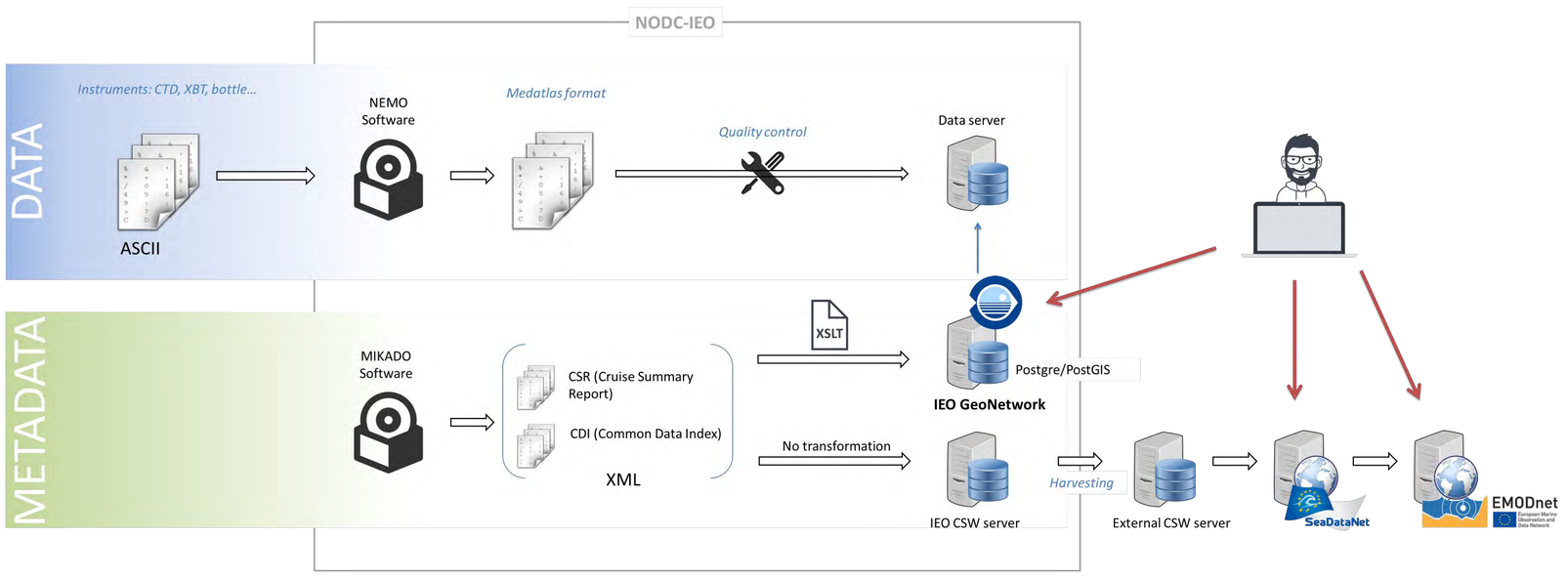

Fig. 1: Scheme of the data and metadata processing carried out by the NODC-IEO. Data to be shared with the SeaDataNet program (CTD, XBT and water samples among others) are formatted and quality-controlled before storing in the server. The associated metadata (CDI) are: i) created using MIKADO, ii) stored in a dedicated server and iii) directly served through the SeaDataNet infraestructure via harvesting. Similar process is done to serve the information about cruises (CSR). In parallel, these metadata are transformed using XSL technology to facilitate their integration into GeoNetwork. Thus, the end-user, can explore and retrieve data through EMODNET, SeaDataNet or directly at the IEO portal.

METHODS

Marine data received at the NODC of the IEO are transformed in ASCII Medatlas format using the NEMO javatool [2]. Data Quality Control procedures are performed to detect missing mandatory information, errors made during the transfer or reformatting, duplicates and outliers. A quality flag is attached to each numerical value in order to not modify raw data. Each dataset is accompanied of a MIKADO [3] XML file, which includes metadata describing the dataset. Subsequently, metadata are shared with the SeaDataNet infrastructure. To properly accomplish with ISO 19139 and INSPIRE Directive (2007/2/EC) and to facilitate the data discovery to the Spanish community, a XSL transformation is applied to metadata. Transformed metadata are distributed through a customized GeoNetwork portal, focused on cataloguing a variety of data served by the IEO.

Why to transform MIKADO XML files?

Cruise Summary Report (CSR) are the usual means for reporting on cruises or field experiments, whereas Common Data Index (CDI) are for reporting on datasets obtained at sea. Usually, CSR and CDI metadata are built using MIKADO software under the requirements of SeaDataNet (http://www.seadatanet.org/). However, these metadata are not adequate to be directly incorporated into a GeoNetwork catalogue. Here, a XSL transformation is applied: to remove SeaDataNet tags that cause failure during importation, to insert additional information and keywords that facilitate data discovery, to add descriptive images and to link with download services, among others. The XSLT is freely available on request.

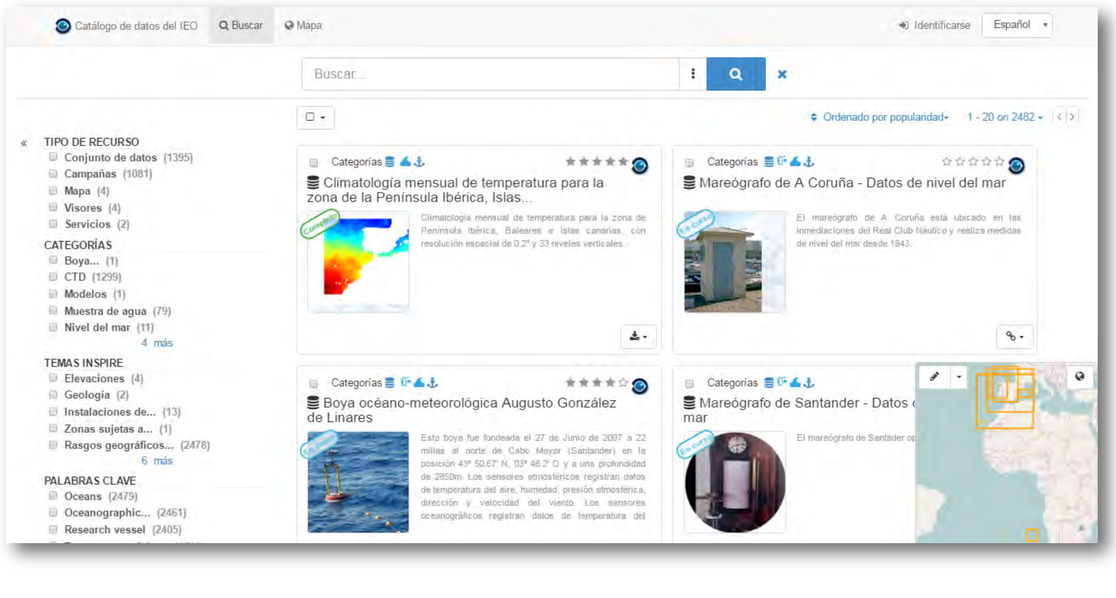

Fig. 2: Search page of the IEO catalogue. As can be observed in the left panel, the catalogue has been configured to include classification by INSPIRE topic and also by type of oceanographic dataset (CTD, buoy, sea level, model…). Metadata are uploaded following the ISO19139 standard with multilingual support (English and Spanish are included).

CONCLUSIONS

The IEO Data Portal is the starting place to find spatial datasets, particularly focused on the Spanish community. The portal facilitates the discovery, multiple usage and dissemination of spatial data.

STANDARDIZATION
Check ISO19139 standard and ensure compliance with the INSPIRE Regulations (as needed). Include CSR and CDI also in your catalogue: XSLT makes the job. Ensure interoperability.

GO

COMMUNICATION
Include non-expert users in your target public . Use native language in both User Interface and metadata (ISO 19139 multilingual). In a first level access, skip technical or confusing words.

CLASSIFICATION
Facilitate finding data by INSPIRE theme, type of oceanographic instrument and even categories related to the Marine Strategy Framework Directive. Include more keywords in English and foreign languages as needed.

REFERENCES
[1] Otero, P., M. Ruiz-Villarreal, G. González-Nuevo, and J.M. Cabanas. 2014. Operational oceanography, end users, and social network sites: An exploratory analysis. Oceanography 27(3):138–147, http://dx.doi.org/10.5670/oceanog.2014.63.
[2] NEMO User Manual v.1.6.3. accessed on 31 Aug 2016 at http://www.seadatanet.org/content/download/25784/176636/file/sdn_Nemo_UserManual_V1.6.3.pdf
[3] MIKADO User manual v.3.3.4 accessed on 31 Aug 2016 at http://www.seadatanet.org/content/download/20278/140561/file/sdn_Mikado_UserManual_V3.3.4.pdf

This work has been supported by EMODNET-DIGEST Project (EASME/EMFF/2015/011: Ingestion and safe-keeping of marine data).

DOWNLOAD THIS POSTER