New SOCIB Data Catalog REST API

Juan Gabriel Fernández, SOCIB (Spain), jfernandez@socib.es
Paz Rotllán, SOCIB (Spain), protllan@socib.es
Cristian Muñoz, SOCIB (Spain), cmunoz@socib.es
Inmaculada Ruíz, SOCIB (Spain), iruiz@socib.es
Miguel Charcos, SOCIB (Spain), mcharcos@socib.es
Miquel Àngel Rújula, SOCIB (Spain), mrujula@socib.es
Xisco Notario, SOCIB (Spain), xnotario@socib.es
Sonia Gómara, SOCIB (Spain), sgomara@socib.es
Miquel Gomila, SOCIB (Spain), mgomila@socib.es
Joaquín Tintoré, SOCIB (Spain), jtintore@socib.es

Introduction

API stands for Application Programming Interface. An API is basically a programmatic tool that is used to interact with a system. More specifically, the new SOCIB API offers the possibility of exploring our data catalog and retrieving the data it contains. This API is a REST API, which means that the requests are issued from a usual HTTP client, available with any programming language.

Regarding interoperability and discoverability, one of the key components that data providers should issue are REST APIs. In this line, SOCIB Data Centre has released in 2018 the Data Catalog API, a REST API that is meant to extend and enhance the capabilities of previous one (Data Discovery).

This new API eases the recovering of different data sources (i.e. observational data) coming from a wide range of platforms (e.g. oceanographical buoy, coastal stations, weather stations, autonomous underwater vehicles, drifter buoys...). Next versions will include also models, images, etc... to the list of available data sources.

The new SOCIB API is addressed to two different types of users: the IT (software developer) user and the data scientific user. For the IT expert, the use will be almost trivial. For scientific data experts, with no experience on this kind of tool, some training will be needed.

A success story is the development of the SOCIB Data Catalog web that fully relies on the capacity of this API.

Main features

One of the main new features is the possibility of discovering and retrieving data without temporal restrictions, unlike Data Discovery API, that only lets you retrieve the latest 60 days of data, at most. Also important is that the data can be filtered with new sets of criteria and retrieved with several resampling methods.

Gridded data type is now also supported (currently only for HF Radar data). Another important breakthrough has been achieved introducing the concept of "data products", that let us provide grouped data in almost any way. The API also supports authentication via API keys, so the tracking of users will now be possible.

A netCDF layer has been developed to act as a middleman between netCDF files (raw data format), queries and JSON formatted responses. This layer aggregates, resamples, prepares and formats the data from the netCDF files that match the query criteria.

This API has been built following the Open API 3.0 specification and a Swagger UI tool has been used to implement an interactive documentation. This is a major improvement since the API itself becomes interoperable.

Data model

The SOCIB API enables users to discover quickly and easily which data products are available at SOCIB. Data products are combinations of data sources grouped by a certain criteria (i.e. all belong to the same campaign or project).

At the core of this API a data model has been developed in order to support this functionality. The main concept in this data model is the so-called data-source, which represents the data produced during the deployment of an instrument. Also a dataset is used to store the data derived from several instruments of a platform (i.e. platforms products such as sea level). The following diagram represents this model in detail:
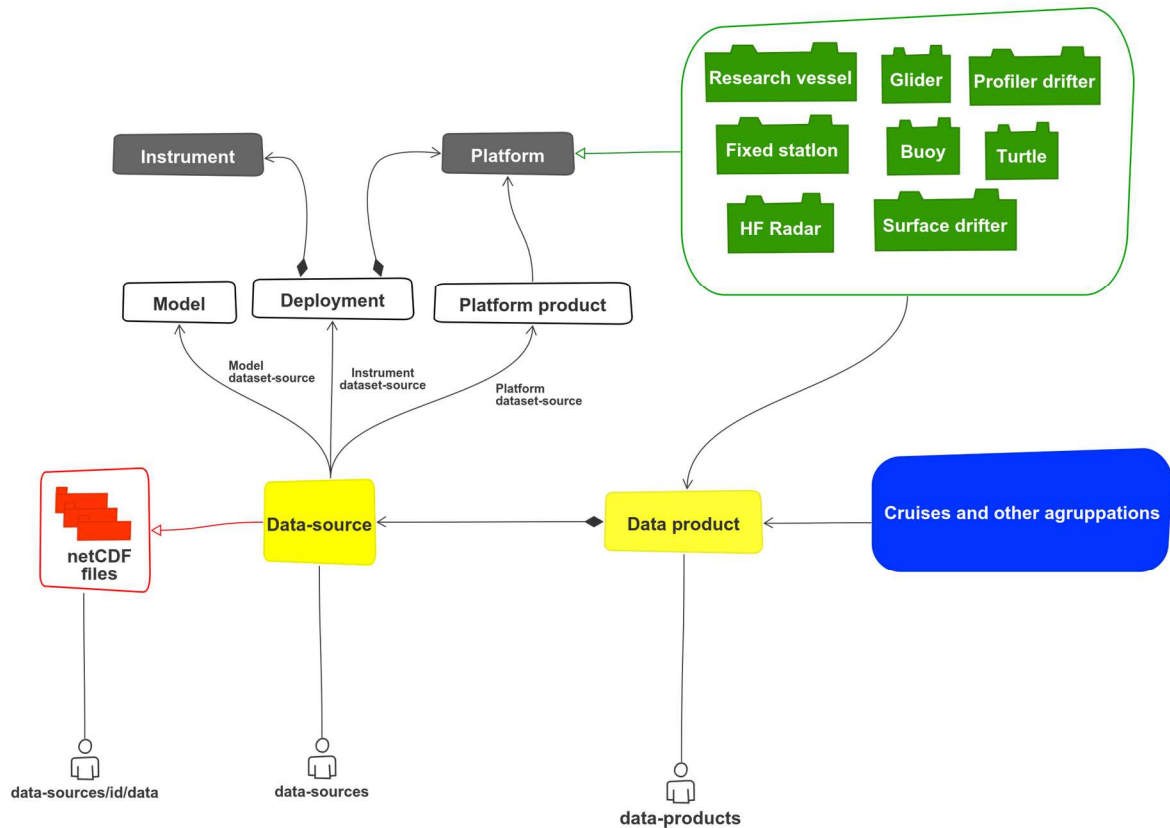

Figure 1: new SOCIB REST API data model

Conclusions

Providing up-to-date tools for users to fetch data from data providers is a must regarding user uptake and engagement. Since the release of the first SOCIB API (Data Discovery), new technologies and methodologies arisen and led to the necessity of readdressing such first approach. The new SOCIB API, the Data Catalog API, is meant to fulfil late requirements in terms of web services and interoperability.