Dealing with (historical) data and making it accessible: Data Inventory and Tracking System (DITS) applied in the scope of the *"4 decades of Belgian marine monitoring"* project (4DEMON)

Marielle Adam, Royal Belgian Institute of Natural Sciences (Belgium), madam@naturalsciences.be
Ruth Lagring, Royal Belgian Institute of Natural Sciences (Belgium), ruth.lagring@naturalsciences.be
Yvan Stojanov, Royal Belgian Institute of Natural Sciences (Belgium), ystojanov@naturalsciences.be
Francis Strobbe, Royal Belgian Institute of Natural Sciences (Belgium), fstrobbe@naturalsciences.be
Thomas Vandenberghe, Royal Belgian Institute of Natural Sciences (Belgium), tvandenberghe@naturalsciences.be

Long-term datasets are important to study global trends in the environment and to define a reference status of an ecosystem. Results of analyses on such datasets allow policy makers to define proper strategies to reach regulatory objectives. Access to these data is also a crucial issue for the scientific community.

Belgium has a long oceanographic data collection history. In that context, the Belgian Marine Data Center, with its continuously growing experience in data rescuing and management, worked out a tool to identify the marine datasets and retrieve as much data as possible.

"4 decades of Belgian marine monitoring: Uplifting historical data to today's needs" (4DEMON)

The tool was developed in the scope of the project *"4 decades of Belgian marine monitoring: Uplifting historical data to today's needs"* (4DEMON). This project aims to build quality-checked, intercalibrated and integrated data on contamination, eutrophication and ocean acidification in the Belgian Continental Shelf (BCS) since 1970. Historical data is being compiled and integrated with more recent data resulting in robust long-term datasets that are made publically available. This will not only safeguard the historic Belgian marine data, but also uplift its value as it can serve many new and cross-disciplinary research objectives and help policy makers to take adequate management actions.

DITS in practice: from data inventory to database and data dissemination

Inherent to these kind of projects on historical data, there are several issues, such as complexity and heterogeneity of the data, potential missing metadata, numerous data formats and data sources. To face these matters, a well-coordinated system to track and to disseminate data needed therefore to be set up.

Accordingly, the "Data Inventory and Tracking System" (DITS) was developed for inventory of (historic) files or datasources like technical reports, logbooks and scientific publications compiled during the project. The submitted datasources, which can be fully described with all required

metadata (like project, data type and data originator), can be combined into 'ingestion datasets' or logical groups of datasources. Original and processing files can be uploaded or refered to and each step of the process can be tracked (Fig. 1). Furthermore, missing information like sampling date or location can be indicated. The tool was designed by and for datamanagers to process data that need further handling, like import in central databases. Data submitters on the other hand can follow-up the status of their data and make updates via the online interface where, all records can be explored and exported.

While DITS serves as a shared tool for data management and traceability, it also allows to centralise, integrate and valorise data compiled in the frame of specific research projects and monitoring programs, forming an important reference for data managers, policy makers and scientists.
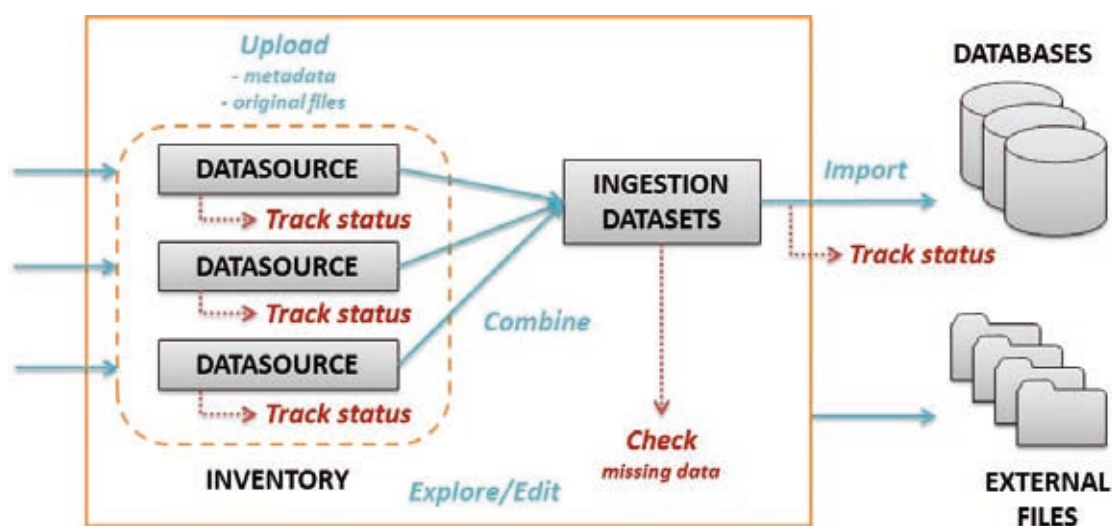

Fig. 1 - Data flow.