'The 'PEGASEAS' project was selected under the European cross-border cooperation programme INTERREG IV A France (Channel) – England, funded by the ERDF."

Promoting Effective Governance of the Channel Ecosystem
Promouvoir une gouvernance efficace de l'écosystème de la Manche

01

Integration and sharing of data on marine ecosystems.

ABSTRACT

Integration and data sharing on marine ecosystem constitute two major contributors to ecosystem management. In this report, we present the different methods that have been used to integrate and share data in the context of Interreg IVA France (Channel) – England projects. Identification of limits and barriers encountered within these projects may help to improve our capacity to lead future projects, notably those involved in marine ecosystem management.

KEY WORDS

ACCESSIBILITY
DATA SHARING
QUALITY OF WORK
REPORTS
WEBSITES

DESCRIPTION OF KEY FINDINGS

Data integration - i.e. how databases are stored

Within Interreg IVA projects such as CAMIS, ChanneLIS, CHARM 3, CRESH, LiCCo, Marinexus, MERiFIC, OFELIA, PANACHE, SETARMS and VALMER, marine ecosystem data were mostly obtained by:

- consultation of pre-existing documents (e.g. legislation, reports, datasets, models outputs, etc.)
- *in situ* samplings (e.g. scientific survey campaigns on a boat in the CHARM 3 project or monitoring techniques using FerryBox, cross-channel transects, Continuous Plankton Recorder, fixed stations within the Marinexus project, seabed towed video , marine birds studies and multi-beam sonar as in PANACHE project)
- molecular studies (e.g., barcoding within Marinexus)
- model outputs (e.g. MARXAN model in PANACHE project)
- satellite imagery.
- direct requests to experts.

In developing the outputs/deliverables within Interreg IVA projects, various methods were used in order to integrate marine ecosystem data. The vast majority of outputs/ deliverables were logically integrated within technical or scientific reports (see Figure 1), as these were a required deliverable of the Interreg projects. Other

integration methods were used including: databases, meetings (oral presentation/ posters), maps, scientific journals papers, conference abstracts, mathematical models, technical reports, web atlas, workshops, websites, etc., however this was to a lesser extent when compared to the examination of the reports (see figure 1).

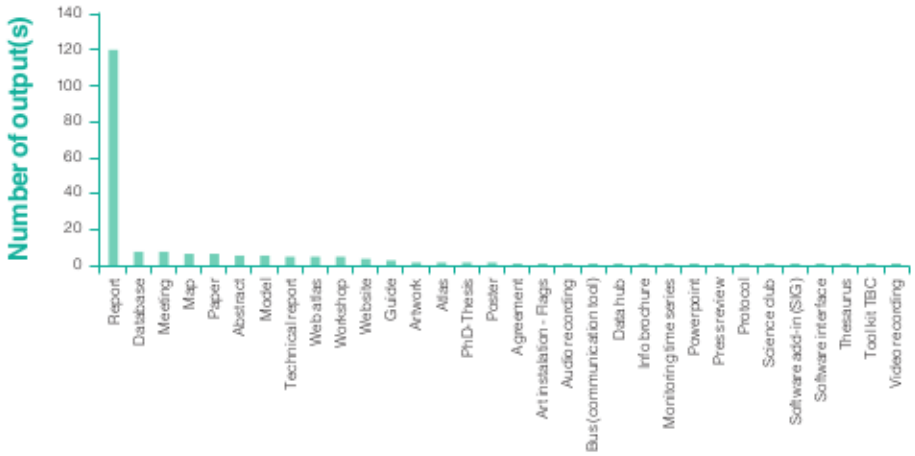

Figure 1. The use of different integration methods within the overall outputs of Interreg IVA projects (across the six PEGASEAS themes).

Data sharing

Methods

Within Interreg IVA projects, raw marine ecosystem data has been shared between partners by email, and are usually not made public.

Unlike raw data, outputs/deliverables of Interreg IVA projects have been made available to the public. For example, the CHARM 3 project provided several maps of species distribution available in the French data infrastructure for marine environment 'SEXTANT'[1]. In addition, several websites have been created in order to facilitate data and information sharing (e.g. CHARM 3[2], PANACHE[3] and Marinexus[4]). Another sharing method was the CHARM 2 Atlas[5]. It was designed in order to (i) integrate diverse marine environment and biological data on habitats, important species and marine living resources in the Eastern Channel, (ii) develop tools to aid decision-making and marine environmental planning, (iii) evaluate and compare cross-border policies and legal frameworks for marine resource management, and (iv) disseminate the Atlas-based information to increase public awareness.

Limits/barriers

Within Interreg IVA projects, raw data have not been made public. Despite the INSPIRE Directive[6], this is a common situation in the domain of information control and knowledge management, intellectual property, data ownership, sensitive data (exploited or threatened species), etc. As such, data integration/sharing is often limited with scientific and industrial sectors becoming increasingly competitive. In fact, nowadays, scientific and industrial sectors are highly competitive. The control of information is therefore important for scientific, industrial or cultural notoriety of the various institutions (universities, research institutes and other organisations involved in research and innovation). Confidentiality and copyright is an essential component of the protection of knowledge: it is the only way to protect know-how and patentable inventions. Due to the enforcement of data systems property, data integration becomes limited and therefore slows down the progress of numerous domains, for example science and governance.

Limited data sharing will inevitably reduce the homogeneity of systems used by partners and therefore multiplies efforts and costs. This strategy forces each user to find the information he/she needs by himself/herself, for example by consulting institutions belonging to other networks than the working group itself. This could

1 CHARM 2 and 3 (undated). *Metadata catalog of spatial data sets*. Available at: http://www.ifremer.fr/sextant/fr/web/charm/geocatalogue

2 CHARM – see: http://www.charm-project.org/fr/

3 PANACHE – see: http://www.panache.eu.com

4 Marinexus – see: http://www.marinexus.org/?lang=fr

5 CHARM 2 (2009): Channel Habitat Atlas for Marine Resource Management. Available at: http://archimer.ifremer.fr/doc/2009/rapport-7377.pdf

6 European Commission (2007). *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*. More information available at: http://inspire.ec.europa.eu/

lead to several biases (e.g., two individuals working on the same theme in a given ecosystem may have different datasets such as, for example different sea surface temperature data extracted from model outputs and satellites). In this example, when data are inaccessible to external institutions (e.g. for physico-chemical parameters), lack of data sharing can then generate differences in the results of scientific studies.

Despite data becoming increasingly reliable, scientists may often spend time checking its reliability due to this wide range of data sources. In that case, the risk is the collection and focus on poor quality data which may generate wrong conclusions. Several project leaders and partners agreed that it would be useful if future project funders will need to make database completion mandatory and in a compatible format (standard). This could become a mandatory deliverable for future projects. Several national and international databases already exist and are particularly efficient (e.g. landings data by the International Council for the Exploration of the Sea).

In contrast to raw data, outputs/deliverables within Interreg IVA projects were accessible (to citizens, scientists, stakeholders, etc.). Nevertheless, only 40% of these outputs were directly accessible (e.g. direct access to a given website, report, etc.) and this implies that a request for desired information needs to be sent to specific persons, especially for reports. The difficulty of data access is accentuated by the fact that it is sometimes problematic to identify which person must be contacted to collect information (or data). This means that stakeholders such as professional organisations, businesses, associations, consultancies and also citizens do not have easy access to project deliverables although technically they are publicly accessible. It may therefore, be worth recommending the inclusion of contact details from which to obtain data as a standard.

Another barrier has been identified by project leaders: data sharing by using websites presents the constraint that they must be regularly updated. The persons involved in a specific project do not necessarily have time (because of fixed-term contracts or other projects in progress) to update these websites and/or format the data to make them compatible to all users. As a result, the websites may become useless if the database or the retained information is out of date. Nowadays, the tendency is that each project has its own website although themes can be sometimes quite similar between projects. The consequence is that the information about a given theme/problem is dispersed and thus its' access is time-consuming. To solve this issue, for a given theme/problem, the use of a generic website (i.e., one which contains several project websites that concern the given theme/problem) or common database (e.g. the Atlantic North East Database accessible through the OSPAR, PANACHE or MAIA website) could be useful, most notably in order to bring together all available information about a theme, rather than scatter it.

Finally, it was noted that socio-economic outputs were more difficult to access than scientific outputs, probably because of their sensitivity (see Figure 2).

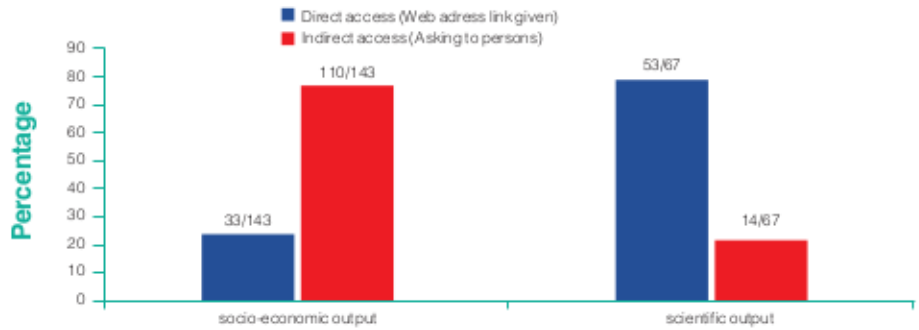

Figure 2. Accessibility to different outputs types within Interreg IVA projects (across the six PEGASEAS themes).

CONCLUSIONS/WORK LEADS

- Raw data are essentially exchanged between experts by email. These data are most often inaccessible to the general public.
- Project outputs are essentially exchanged by reports. Obtaining these reports often requires personal request to producers.
- Limits and barriers mainly concern sharing methods.
- Where the data are public and accessible, data are, often not easily found (direct vs. indirect access).
- Without the enforcement of a consistent data system, data integration/sharing becomes limited (risk of loss of quality of information/data).
- Data sharing via websites involves making regular updates.
- There may be multiple websites within a given theme.
- Scientific publications are not always publicly available.
- Socio-economic outputs were more difficult to access than scientific outputs.
- Data access is still too limited for public stakeholders: with no data available, the interpretation of results is difficult for decision-makers.

Recommendations

- It is important to produce common methodologies for the acquisition, storage, classification and validation of data (and metadata) to reduce time delays and to enhance the quality of work.
- A solution to make data sharing more efficient would be for funders to require all project partners to fill in databases, that are using a compatible format (i.e. usable by all). The implementation of a generic website (i.e., one which integrates several project websites that deal with the given theme/problem) could also be relevant.
- Outputs/deliverables must be attributed to the author, with a key person as a contact.

Authors: Auber Arnaud (Ifremer), Delpech Jean-Paul (Ifremer)