

Mediterranean Marine Science

Vol 12, No 3 (2011)

Vol 12, No 3 (2011) special issue



Safeguarding, Integrating and Disseminating Knowledge on Exploited Marine Ecosystems: the Ecoscope

J. BARDE, P. CAUQUIL, P. CHAVANCE, P. CURY

doi: [10.12681/mms.69](https://doi.org/10.12681/mms.69)

To cite this article:

BARDE, J., CAUQUIL, P., CHAVANCE, P., & CURY, P. (2011). Safeguarding, Integrating and Disseminating Knowledge on Exploited Marine Ecosystems: the Ecoscope. *Mediterranean Marine Science*, 12(3), 45–52.
<https://doi.org/10.12681/mms.69>

Safeguarding, Integrating and Disseminating Knowledge on Exploited Marine Ecosystems: the Ecoscope

J. BARDE, P. CAUQUIL, P. CHAVANCE and P. CURY

IRD, Centre de Recherche Halieutique Méditerranéenne et Tropicale (CRHMT),
Avenue J. Monnet, BP 171, 34203 Sète, France

Corresponding author: julien.barde@ird.fr

Abstract

Ecosystem Approach to Fisheries (EAF) is an application of sustainable development concepts to exploited marine resources management. In particular, this approach aims to improve decision making by facilitating information and knowledge sharing among the different stakeholders in the domain. The CRHMT in Sète studies Mediterranean and tropical fisheries by taking into account the related ecosystems and their elements. Informational resources (IR) needed for an ecosystem approach to manage these fisheries are collected at a global scale: from local databases as well as from several sources distributed around the world. While Internet enables IR sharing, IR discovery still remains complicated because of their heterogeneity (from a thematic, technical and semantic point of view). The challenge for IR management in this domain, like in many others, is to set up an information system which facilitates (online) IR discovery, access and, if possible, certain treatments (like a simple subset or aggregating data from one or different sources...) through dedicated portals. The use of standards is crucial for such systems interoperability. Indeed, standards facilitate external users understanding of what kinds of data are available in the different sources and how they can get them. We present an information system which aims to satisfy these needs by complying with the main standards of the domain (for metadata, ontologies, knowledge and spatial information management). Attention will be paid on using a generic approach to set up an architecture which can be reused in other case studies.

Keywords: Ecosystem Approach to Fisheries; Information systems; Metadata; Ontologies; Knowledge base; Interoperability; Databases.

Introduction

The informational resources¹ (IR) required

1 any kind of data, information, knowledge produced (including treatments) to study a domain (regardless of the format). According to the domain and users: an observation, a report, a map, a picture, a video, a dataset,

to set up an efficient management of environmental domain are by nature, heterogeneous, like the different elements interacting in this domain. Moreover, the studying of ecosystems under human pressure, like oceans, requires the collection of IR from a wide spectrum of disci-

a database, a model...

plines, from human sciences to environmental sciences. Since Reykjavik conference on responsible fisheries, the Ecosystem Approach to Fisheries (EAF) is promoted by international organizations to set up an alternative management approach which leans towards the integration of the various information acquired on the main elements of an ecosystem (natural or anthropic elements) by the different stakeholders in the domain. In practical terms EAF can be considered as an application of sustainable development to fisheries management just as ICZM for coastal management. FAO's definition of EAF is the following in [FAO]:

“This approach strives to balance diverse societal objectives, by taking into account the knowledge and uncertainties about biotic, abiotic, and human components of ecosystems and their interactions and applying an integrated approach to fisheries within ecologically meaningful boundaries”.

The success of the underlying IR (knowledge and uncertainties) management leans on the integration of both technical and thematic constraints. From a technical point of view, metadata and ontologies are needed components of any information system to ensure an efficient management of heterogeneous IR. Indeed, the ability to locate existing IR requires the implementation of standards to manage human and machine-understandable descriptions at a worldwide scale. Standardized metadata sheets (and related catalogs) enable a syntactic interoperability where controlled vocabularies (thesaurus, ontology...) enable semantic interoperability by taking into account indexation quality issues (synonyms, languages...). Metadata and ontologies are major components to drive the information system during IR discovery (to improve the results relevance for a given query: for example by expanding or refining it thereafter) and access processes. Nevertheless, from a thematic point of view, the setting up of a domain ontology requires an efficient description of the elements and their interactions in the ecosystem by related experts. It is thereafter possible to set up a knowledge base. In this paper, we present an overview of the approach used to set up

the information system's architecture which aims to satisfy these needs. In section 2, we describe the generic and specific (at the CRHMT) contexts and the related issues in terms of IR heterogeneity management. In section 3, we describe how we plan to improve IR management by implementing helpful standards: the key point of IR discovery (through a proper metadata and ontologies management), the improvement of IR access and knowledge sharing. In section 4, we give some details on the implementation choices we made (standards, softwares...) to set up our system and illustrate some results (GUI snapshots).

Context and Issues

This section summarizes the international and local scientific contexts from a thematic and technical point of view, the issues as well as the main goals of this project.

International context

IR heterogeneity and sharing issues are actually well known and described in different domains and institutes. This is due to historical reasons: the succession of projects and related tools based on the ongoing technologies (at different periods). However, with networks, especially Internet, interoperability between distributed systems is made possible by implementing standards for metadata, IR and their treatments (NOGUERAS-ISO *et al.*, 2004). Currently, depending on domains and IR kinds, there is a broad choice of standards. Because of W3C recommendations, most of these standards use XML (and related XML schemas) as a support of interoperability to share metadata and data. Most of the time, some tools to implement them properly already exist. It is thus needed, for the previous information systems to safeguard and share their IR, to take these standards into account either by integrating new components to implement these standards or by building a new information system on top of them. In the environmental and marine domains, metadata and ontologies importance is now well recognized. The common work of ISO TC211 [ISO] and

OGC [OGC] brings some efficient and widely used standards for spatial metadata, data and (Web) Services management. European union considers some of them are essentials so that European countries have to implement them (INSPIRE European directive [INSPIRE]). At a global scale, the scientific community involved in IR sharing in the marine domain encourages as well stakeholders to implement these standards (MILLARD *et al.*, 2005; GRAYBEAL *et al.*, 2003; IODE). Regarding knowledge management, W3C suggests to use a set of standards related to its Semantic Web activity (W3C, 2004). Some major projects based on Semantic Web standards already exist to manage knowledge related to marine domain. The first task of setting up a new information system architecture consists in choosing appropriate standards to satisfy both its users needs as well as community expectations in terms of interoperability.

Local context and main goals of Ecoscope project

The Mediterranean and Tropical Halieutic Research Center (CRHMT), created in 2001 in Sète (France) is a joint research structure partnered by Ifremer, IRD and the University of Montpellier II. CRHMT is focusing its research on studying certain tropical and Mediterranean fisheries and the related marine ecosystems. This center develops multidisciplinary integrated research on EAF domain within the context of global climatic change and overexploitation. Studying these elements requires different kinds of IR, so that IR management represents an important part of CRHMT research effort. Moreover, as the global scale of the large pelagic fishes distribution makes it necessary to collect data from all over the world (which means from heterogeneous data sources remaining on different technologies), the international context in IR management brings some generic solutions which have to be taken into account. The CRHMT thus set up the Ecoscope project² whose main idea, first formalized by Ulanow-

icz (ULANOWICZ *et al.*, 1993) and taken back by Cury (CURY, 2004), consists in enabling a reasoned management of knowledge related to these fisheries by using an ecosystemic approach. In practical terms, this project aims to sustain EAF by improving IR management (descriptions, searching, access, treatments) to support decision making process. In particular, because of the IR heterogeneity, the prior goals are to archive, articulate and facilitate access to the knowledge acquired by different past, ongoing or future projects. To satisfy these needs, the Ecoscope project set up an original information system which integrates different components. A component for metadata management is needed but not sufficient to guarantee the indexation quality of IR in the catalog. A knowledge management component is thus set up in order to improve indexing and thereafter the searching process (by expanding or refining queries). Thereafter, Ecoscope aims to facilitate physical access to the IR described in the catalog through generic methods.

Improving IR discovery and access in EAF domain

The prior goal of Ecoscope project is to enable IR discovery for all (local and external) users in order to answer to simple queries like: "Find all databases containing data on Swordfish in Indian Ocean during the following period..." To get results from our application, we have to set up a standardized catalog containing the descriptions (metadata) of existing IR (see section 3.1) formalized in a machine understandable way: from a syntactic point of view (metadata elements sets) as well as from a semantic point of view (ontologies). Once aware of existing IR, their accessibility is facilitated thereafter.

Kinds of IR at CRHMT

In practical terms, according to the different projects and related users, CRHMT needs the following kinds of IR:

- from a thematic point of view: IR are related to ecosystems and their different biotic, abiotic,

2 www.ecoscope.org

and human components. They deal with species observations (biological sampling, biometry, tracking...), trophic networks, environmental data (wind, temperature, SST, chlorophyll, sediments...) designated by different terms (semantic heterogeneity: terms, languages...),

- from a technical point of view: different kinds of IR (non or semi-structured data, databases, spatial information, treatments: spatial analysis, quality controls..., models, pictures, videos...) which are most of the time managed in different systems (softwares, versions) with different formats (SQL, XML, vector: shp, SFS, GML, KML..., raster data: O&M, NetCDF...).

However, it is difficult for users to be aware of existing IR, even for local users (which means nearly impossible for external stakeholders). Ecoscope aims to fill the lack for such a tool. CRHMT aims now to homogenize its IR management to share them with its partners through a single meta-portal built on top of existing ones.

Relevance of metadata standards

We consider that a metadata standard instance (or metadata sheet) is just a way to make people aware of existing IR and it could be done with numerous metadata standards. Indeed, so far, there is not a single (magic) metadata standard to describe accurately any kind of IR (picture, spatial information, species observation...). However, the Dublin Core Metadata Initiative (DCMI) [DCMI] has been created to summarize the content of any kind of IR by using a light and very generic metadata elements set. Thus DCMI can help to bridge the gap between heterogeneous (domain specific) metadata standards. Nevertheless, it's clear that, according to IR kinds, some of them should be described as well in a better way by using dedicated standards : like Darwin core or ABCD for species Observations, or SensorML for sensors data (CTD...) on different kinds of Platforms (buoy, satellites, vessels...). Figure 1 illustrates how all these different metadata standards answer specific needs and all share core metadata elements which match the DCMI metadata elements.

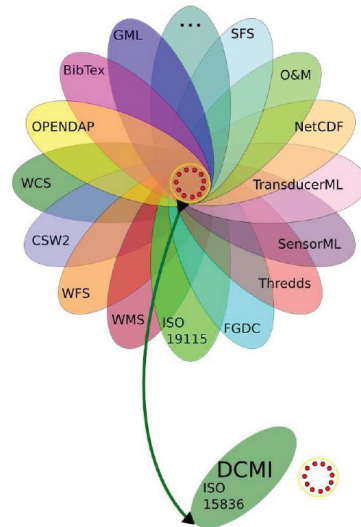


Fig. 1: Different metadata elements: same DCMI core.

As most of the IR we aim to describe are kinds of spatial information, we decided to focus first on describing them with ISO 19115/39 in a simple way (by using mainly the core metadata elements which basically match the DCMI elements). Moreover, to ensure the interoperability with similar catalogs (expansion of requests / remote harvesting) of our partners, regardless of standards and tools chosen to implement them, we need to implement the CSW standard. This way, we are compliant with ISO, OGC and W3C recommendations as well as the INSPIRE directive and can use existing tools to implement these complicated standards. However we keep in mind that we should in the short term be able to implement any other relevant standard for our domain. This is the reason why we wanted to manage metadata with a multi-metadata standard software (see section 4) to make it possible in the long term. But, regardless of the metadata standards, cataloging IR in a standardized way from a syntactic (XML schemas) point of view is not enough without controlling XML tags content: a common semantic referential is also needed to facilitate their discovery.

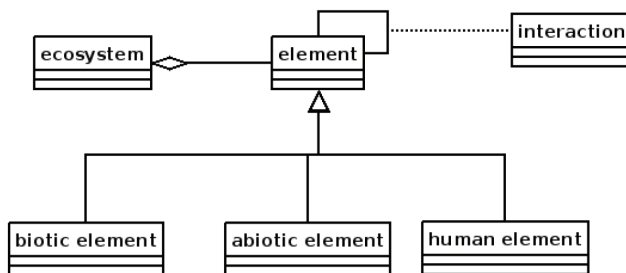


Fig. 2: FAO's EAF definition as an UML class diagram.

Necessity of ontologies for IR discovery and knowledge management

This is another essential component of the Ecoscope project. We aim to describe in an homogeneous way different catalogs by using a homogeneous semantic. This is crucial to guarantee query results quality. Indeed, if different catalogs use different vocabularies and are requested by using CSW standard, how can we match in an automated (which means machine understandable) way the content of their metadata sheets without ontologies (synonyms, languages)? This is especially true for the (XML) tags used to describe the thematic scope as well as the spatial extent of the IR. Either catalogs share the same semantic and spatial referential or it's complicated to query them on the fly through single portals. The semantic management allows users to get answers from different systems through a single query by using ontologies to match relevant results in different sources (related to similar concepts even if indexing is based on different terms). Unfortunately, so far, none ontology on EAF domain exists. A challenge for this task is to set up our own ontology by complying with other conceptsschemes. Indeed, ongoing efforts to build ontologies are increasing [4, 7] but difficult to integrate at a local point of view [Pettman et al, 2006]. To fill this lack, our first task has been to suggest a pattern which can be used to inventory the knowledge owned by the CRHMT on marine ecosystems and their elements. We thus translated the FAO

EAF definition as an UML class diagram³ (Fig. 2). This diagram has been used to generate our SKOS (RDF/XML serialization) thesaurus and thereafter the OWL ontology for richer semantic relationships (ongoing work). The thesaurus is needed to homogenize indexations between our network of catalogs (and thereafter enable queries expansion or refining). The ontology with richer semantic relationships will be use to set up a knowl-edge base with fact sheets, networks representations. Being able to retrieve existing IR is obviously the preliminary step to access and treat them thereafter.

Improving IR treatment

Once users have located some relevant IR for their work through metadata, the next step and main issue remains in delivering the related IR in one (or several) generic format(s) understandable by anyone (with appropriate syntactic and semantic formats). DBMS are needed to guarantee a good IR management but we consider that RDBMS heterogeneity is useless, at least in the case of a small organizations with limited human resources. Thus, we aim to migrate existing databases at CRHMT from old or heterogeneous DBMS towards a single one which is efficient and standardized (and OpenSource to avoid software license fees, especially for developing countries). Attention will be paid on homogenizing Conceptual Data Models (CDM) while

³ As this formalization is very generic, it could be applied to other domains dealing with the exploitation of natural resources.

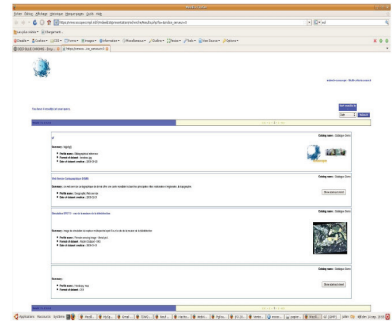
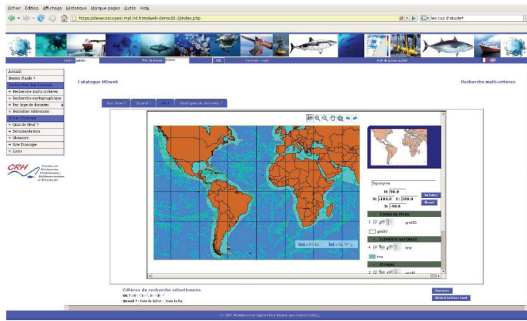


Fig. 3: Searching engine GUI to query different metadata catalogs.

databases migration by sharing referential for spatial information and taxonomy. Moreover, it is needed to keep on being compliant with the main formats for (physical data management). Then IR can be serve in different ways (database connection or Web Services for example).

Implementation choices

To set up this architecture, we have chosen to reuse existing softwares and integrate them as different components of our information system. Each of them is in charge of different tasks. The filling of this system will be done by focusing on different case studies.

A standardized SDI

Our main goal is to reuse OpenSource software components to manage the IR in a proper way, by implementing the relevant standards:

- generic and standardized metadata services to describe them (ISO 19115/39, DCMI, SensorML...) and search (CSW-2) by integrating MdWeb Software which uses Postgres RDBMS and Mapserver,
- standardized formats to manage and access some essential kinds of IR. Spatial information will be managed with the recommended formats managed into Postgis (main OGC formats: SFS for SQL, GML, KML, or O&M for XML as well as other standardized formats like NetCDF...). Knowledge will be managed with the recommended W3C formats (SKOS/RDF/OWL...) by us-

ing the Jena Semantic Web framework (into Postgres RDBMS),

- standardized Web Services to serve and treat them (WMS, WFS, WCS...) by reusing OpenSource software like Mapserver or GIS clients like Qgis, GvSig, Grass...and similar Web Services for knowledge management (by using RDF or OWL formats). Off course, from the user's point of view, we hide these details of the architecture to make it as simple to use as possible.

GUIs set: user's point of view

An homogeneous (regardless of standards) set of friendly GUIs for online services:

- IR discovery: by editing detailed metadata (Fig. 4.2) and consulting them through multi-standard search engine (Fig. 4.2),
- IR management : by homogenizing the semantic (CDMs...) and formats coscope will enable generic treatments: IR subset or aggregation...,
- IR treatment : fact sheets (Fig. 4.2) about ecosystems and their elements (properties: terms, indicators..., and interactions: semantic network, Web Mapping...).

Case studies

Among topics of interest, Ecoscope project is first going to focus its cataloging and managing efforts on the Mediterranean Sea (Gulf of Lion, ..) and Benguela (large) Marine Ecosystems (Fig. 5). In these ecosystems, Ecoscope will first focus on the following elements (biotic,

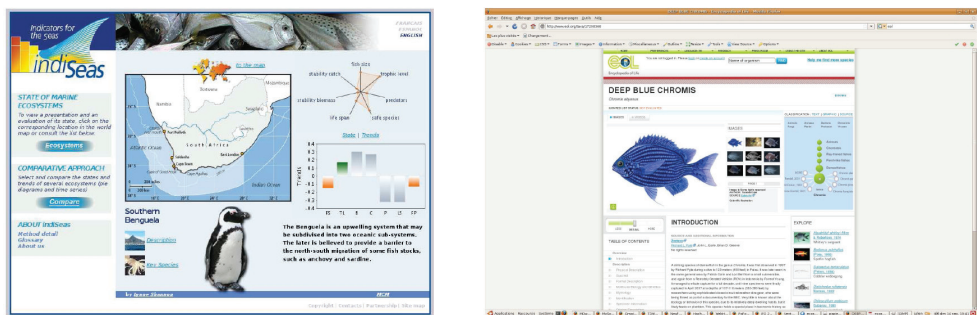


Fig. 4: Example of fact sheet from Indiseas and EOL project.

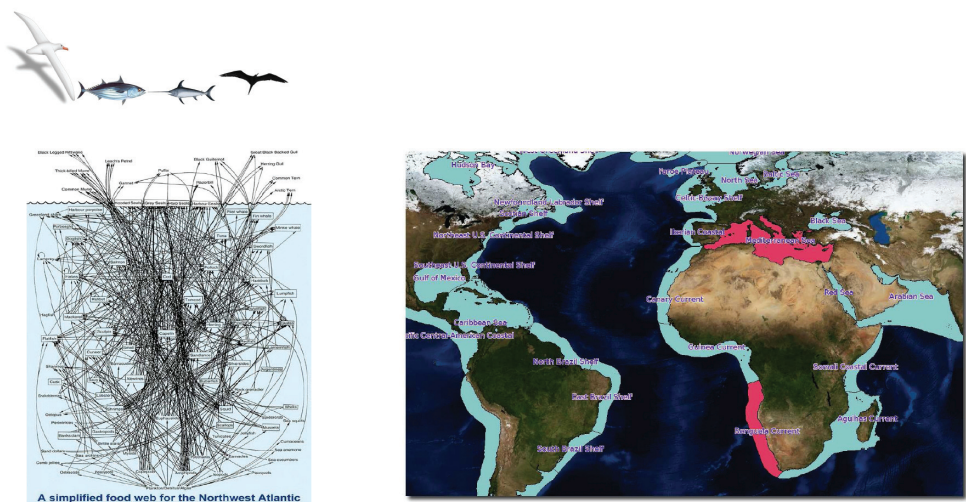


Fig. 5: Case studies for Ecoscope project.

abiotic and human) and interactions: large and small pelagic fishes and related species in the food web (predator-prey relationship) of these fisheries (Fig. 5). It is thus needed to improve IR management on these topics to illustrate the interest of a new IS infrastructure. Because of the generic approach we will expand it to other cases studies in the long term. Among the available kinds of IR at CRHMT, Ecoscope aims to share priorly metadata on the following ones: spatial information, databases and models (by describing both structure through conceptual and physical data models -as UML class diagrams and content: data as SQL, XML, netcdf files or treatments...) and resulting information

and knowledge (models outputs, indicators, references: book, article, reports...), pictures...

Conclusion

As many laboratories or institutes, CRHMT currently faces a lack of IR management tools to satisfy the basic needs of users coming from different domains. Off course, this situation is due to historical reasons (like computer sciences evolution in terms of technologies, formats, networks...) as well as human resources reasons. At a period and in a domain where historical and international data can really change decision making, it is needed to change the current man-

agement of IR quickly. The CRHMT thus set up the Ecoscope project which aims to improve IR sharing by:

- helping users to discover the existing and relevant IR for EAF through better descriptions and searching processes,
- setting up a knowledge base (based on an ontology) to share and aggregate knowledge owned locally on ecosystems and related elements with external stakeholders,
- being compliant with the main standards to manage (ISO/OGC/W3C) and treat (visualization) the IR: interoperability with partners. • focusing on prior case studies:
 - locations: Benguela and Mediterranean as Large Marine Ecosystems,
 - ecosystem elements: small and big pelagic fishes and related preys and fisheries.

Since both Ecoscope's methodology and technology are generic and OpenSource, the approach can be applied and expanded thereafter to other thematic case studies. From a technical point of view, we aim to expand and validate the interest of this approach by deploying similar nodes based on this architecture in different locations. Interoperability is made possible as the different information systems all share the same semantic and spatial information referential to guarantee the efficiency of IR management and the feeding by different nodes. Nodes deployment is going to be easier because of the choices made in term of metadata and ontologies standards as well of tools to implement them. Then, users will then be able to query this network through a single portal to get details about existing IR and eventually access them.

References

CURY, P., 2004. Tuning the ecoscope for the ecosystem approach to fisheries. *Mar.Ecol. Prog.Ser.*, p 272–275.

DCMI, 2005. Dublin core metadata initiative. <http://dublincore.org>.

FAO, 2003. The ecosystem approach to fisheries. FAO Guidelines for Responsible Fisheries, (4):112.

FININ, T., SACHS, J. & PARR, C., 2007. Finding Data, Knowledge, and Answers on the Semantic Web. In *Proceedings of the 20th International FLAIRS Conference*. AAAI Press .

GRAYBEAL, J., BERMUDEZ, L., BOGDEN, P., MILLER, S. & WATSON, S., 2003. Marine metadata interoperability project: Leading to collaboration. Technical report.

ISO/TC 211, 2011. Geographic information/ Geomatics, <http://www.isotc211.org/>

Infrastructure for Spatial Information in Europe, 2011. <http://www.ec-gis.org/inspire/>.

Marine Overlays on Topography for Annex II Valuation and Exploitation. <http://www.iode.org/motiveweb/index.php>

MILLARD, K., 2005. Using xml technology for marine data exchange. A position paper of the marinexml initiative. Technical report, Mars.

NOGUERAS-ISO, J., ZARAZAGA-SORIA, F.J., LACASTA, J., BJAR, R. & MEDRANO, P.R., 2004. Metadata standard interoperability: application in the geographic information domain. *Computers, Environment and Urban Systems*, 28:611–634.

Open Geospatial Consortium, 2011. <http://www.opengeospatial.org/>.

PETTMAN, I., 2006. Knowledge retrieval in aquatic ecology and fisheries: Do we need (and can we afford) ontologies? In *IAMS LIC Conference Proceedings 2005*.

SACHS, J., 2006. Using the Semantic Web to Support Ecoinformatics. In *Proceedings of the AAAI Fall Symposium on the Semantic Web for Collaborative Knowledge Acquisition*. American Associate for Artificial Intelligence.

ULANOWICZ, R.E., 1993. Inventing the ecoscope. In V. Christensen and D. Pauly (eds) *Trophic models of aquatic ecosystems, ICLARM Conference Proceedings* 26:ix-x.

World Wide Web Consortium (W3C), 2004. Web ontology language (owl) use cases and requirements. <http://www.w3.org/TR/webont-req/>.