NeOn-project.org

NeOn: Lifecycle Support for Networked Ontologies

Integrated Project (IST-2005-027595)

Priority: IST-2004-2.4.7 – "Semantic-based knowledge and content systems"

7.2.1 Inventory of Fisheries Resources and Systems

Deliverable Co-ordinator: Yves Jaques

Deliverable Co-ordinating Institution: Food and Agriculture Organization of the United Nations (FAO)

This document contains the results achieved in documenting and detailing a set of fisheries information systems and resources for the purposes of ontology experimentation by NeOn partners, and for the use of the work package seven case study.

Document Identifier:	NEON/2007/D7.2.1/v1.0	Date due:	February 28, 2007
Class Deliverable:	NEON EU-IST-2005-027595	Submission date:	March 30, 2007
Project start date:	March 1, 2006	Version:	v1.0
Project duration:	4 years	State:	Final
		Distribution:	Restricted

NeOn Consortium

This document is part of a research project funded by the IST Programme of the Commission of the European Communities, grant number IST-2005-027595. The following partners are involved in the project:

Open University (OU) – Coordinator Knowledge Media Institute – KMi Berrill Building, Walton Hall Milton Keynes, MK7 6AA United Kingdom Contact person: Martin Dzbor, Enrico Motta E-mail address: {m.dzbor, e.motta} @open.ac.uk	**Universität Karlsruhe – TH (UKARL)** Institut für Angewandte Informatik und Formale Beschreibungsverfahren – AIFB Englerstrasse 28 D-76128 Karlsruhe, Germany Contact person: Peter Haase E-mail address: pha@aifb.uni-karlsruhe.de
Universidad Politécnica de Madrid (UPM) Campus de Montegancedo 28660 Boadilla del Monte Spain Contact person: Asunción Gómez Pérez E-mail address: asun@fi.upm.es	**Software AG (SAG)** Uhlandstrasse 12 64297 Darmstadt Germany Contact person: Walter Waterfeld E-mail address: walter.waterfeld@softwareag.com
Intelligent Software Components S.A. (ISOCO) Calle de Pedro de Valdivia 10 28006 Madrid Spain Contact person: Richard Benjamins E-mail address: rbenjamins@isoco.com	**Institut 'Jožef Stefan' (JSI)** Jamova 39 SI-1000 Ljubljana Slovenia Contact person: Marko Grobelnik E-mail address: marko.grobelnik@ijs.si
Institut National de Recherche en Informatique et en Automatique (INRIA) ZIRST – 655 avenue de l'Europe Montbonnot Saint Martin 38334 Saint-Ismier France Contact person: Jérôme Euzenat E-mail address: jerome.euzenat@inrialpes.fr	**University of Sheffield (USFD)** Dept. of Computer Science Regent Court 211 Portobello street S14DP Sheffield United Kingdom Contact person: Hamish Cunningham E-mail address: hamish@dcs.shef.ac.uk
Universität Koblenz-Landau (UKO-LD) Universitätsstrasse 1 56070 Koblenz Germany Contact person: Steffen Staab E-mail address: staab@uni-koblenz.de	**Consiglio Nazionale delle Ricerche (CNR)** Institute of cognitive sciences and technologies Via S. Martino della Battaglia, 44 - 00185 Roma-Lazio, Italy Contact person: Aldo Gangemi E-mail address: aldo.gangemi@istc.cnr.it
Ontoprise GmbH. (ONTO) Amalienbadstr. 36 (Raumfabrik 29) 76227 Karlsruhe Germany Contact person: Jürgen Angele E-mail address: angele@ontoprise.de	**Asociación Española de Comercio Electrónico (AECE)** C/Alcalde Barnils, Avenida Diagonal 437 08036 Barcelona Spain Contact person: Jose Luis Zimmerman E-mail address: jlzimmerman@fecemd.org
Food and Agriculture Organization of the United Nations (FAO) Viale delle Terme di Caracalla 1 00100 Rome, Italy Contact person: Marta Iglesias E-mail address: marta.iglesias@fao.org	**Atos Origin S.A. (ATOS)** Calle de Albarracín, 25 28037 Madrid Spain Contact person: Tomás Pariente Lobo E-mail address: tomas.parientelobo@atosorigin.com

Work package participants

The following partners have taken an active part in the work leading to the elaboration of this document, even if they might not have directly contributed writing parts of this document:

Universidad Politécnica de Madrid (UPM),

Universität Karlsruhe – TH (UKARL),

Ian Pettman,

Kearon McNicol,

Louise Miles.

Change Log

Version	Date	Amended by	Changes
0.1	22/11	Yves Jaques	Reworked based on partner comments
02	13/02	Yves Jaques	Sections on validation added
0.3	16/02	Yves Jaques	Results completed
0.4	02/03	Yves Jaques	Post review
0.5	02/03	Yves Jaques	Post QA
1.0	23/03	Aneta Tumilowicz	Final QA

Executive Summary

This document describes the work of 7.2.1 Inventory of Fisheries Resources and Systems, the results achieved, and the validation performed in describing and classifying an inventory of information systems that will be considered for inclusion in the Fisheries Stock Depletion Assessment System (FSDAS) knowledge base built upon the NeOn architecture. This knowledge base will function as a decision support system for fisheries assessment experts, helping them to make accurate, consistent and timely fish stock assessments.

The inventory was begun by developing a methodology (see Appendix A) that defined the process and procedures to follow, and included a validation component. Scope was also considered, and included both case study needs and NeOn partner needs.

Interviews with fisheries experts were used to kick-start the inventory process. From these interviews which generated approximately seventy systems, a second group was extrapolated based on similarity of purpose/content to the seed set.

The inventory eventually included 140 systems, stored in a relational database from which 33 were selected for further detailing and 28 finally detailed. For detailed systems, system managers were contacted for information and XML files conformant to a schema based on the inventory methodology were created for each system or federated system.

Each step in the process was followed by a validation exercise. Scope coverage was 76%, largely due to the lack of available systems covering effort data. The percentage of systems finally detailed from the selected set was 85%, due to the difficulty of detailing some systems for which system owners did not respond to our requests for information. The final results for the actually detailed systems showed that the inventory had managed to achieve 86% coverage of the hoped for data.

Although a validation threshold was not included in the methodology, it seems clear from the abundance of detail data (350 pages of XML files) and the broad scope of detailed systems that there is more than enough upon which to base both the case study and any partner experiments.

Apart from the methodology appendix already mentioned, the document contains four annexes; the interviews with fisheries experts, a table detailing the relative weight given to system descriptors for data collection and validation purposes, the XML schema used for system detailing, and a list of acronyms used in the deliverable.

Table of Contents

List of figures

1. Overview

Deliverable 7.2.1 Inventory of Fisheries Resources and Systems contains a methodology, a set of results achieved and a validation exercise in describing and classifying an inventory of information systems that will be considered for inclusion in the Fisheries Stock Depletion Assessment System (FSDAS). FSDAS will be a knowledge base built upon the NeOn architecture that will function as a decision support system for fisheries assessment experts, aiding them to make accurate, consistent and timely fish stock assessments.

The methodology is included as an annex. Also created as part of this deliverable was a database of systems that were considered for inclusion (the basic list of systems is included in the results), as well as a series of detailed data for systems that were finally selected. Neither of these sets were included in the main document as they are not in document-like formats. The data were produced in this way as it was more useful for the purposes of the project, as they can be easily reprocessed using various machine languages such as SQL and XSL.

WP7 urges all partners interested in using the data described by 7.2.1 to have a look at the acceptable use policy found in this document.

1.1 Method

In order for the inventory of fisheries information systems to be consistent, it was backed by a coherent methodology (see Appendix A) that defined inventory processes and procedures to insure that the inventory gathering and the important aspects of every system were captured in a systematic way.

1.2 Scope

The inventory gathering exercise attempted to consider a wide array of possible system types from content management systems, to statistical databases. For a complete list please refer to *Section 1.2, Scope* in the methodology document.

System domain was also considered, as fish stock assessments benefit from data coming from a wide variety of domains, from the physical sciences such as oceanographic data, to the social sciences such as demographic data.

1.3 Results

The deliverable was subject to extremely tight timing, about three months considering that it spanned either side of winter holidays. The methodology was largely completed within the first month, as were the initial interviews with fisheries experts. The overall inventory was begun in December and grew by mid-January to 140 systems, from which 33 were selected for further detailing, and 28 ultimately detailed.

Detailing began in January and continued until the deliverable date. As might be expected, this was the most difficult aspect, as it required in many cases interaction with outside systems managers. As they were not project participants, and had no particular stake in NeOn, their replies were often slow or non-existent. Thus some systems were not detailed as fully as had been hoped for, though information is likely to keep trickling in over the next months.

As of this report, from the 28 systems detailed, 24 can be considered to meet validation requirements (above the 80% validation mean). For those systems not completely meeting validation, this was due to the impossibility of detailing them completely due to the non-interest of some system owners together with scant system documentation.

1.4 Acceptable use policy

The data systems defined and detailed in this deliverable are to be used for the purposes of NeOn work package partner experiments for the scope of the NeOn project, and for the purposes of the NeOn WP7 case study. Any other use of these data systems must be requested specifically from the actual data system owner. FAO makes no claim or warrant as to the suitability of the data for any specific purpose. For FAO data please refer to the FAO copyright statement:

http://www.fao.org/copyright_en.htm

2. Methodology development

The methodology (see Appendix A) was elaborated in the first few weeks of the deliverable. Despite a standard literature search, a satisfactory information system inventory gathering methodology tailored to the needs of the NeOn case study was not located. The decision was taken to apply a standard software engineering approach to the inventory by considering it as a type of information system project.

In this spirit, the methodology was organised first by looking at requirements, both those of the case study and of NeOn partners. This resulted in a purpose section for the methodology, together with a scope for system types and domains. The requirements also naturally led to a set of specific criteria to use in selecting systems, and in selecting a subset for further detailing.

In order to take a user-centric approach to the process, the methodology defined a subtask of interviewing fisheries experts to learn more about the fish stock assessment process and the sources of data used. To help insure some standardisation, an interview template was prepared. This task was used to help in initial system discovery.

The methodology then went on to describe the process to follow in building the initial inventory, a data-gathering activity that was compared to the design phase of an information system engineering process. The criteria culled from the requirements were then applied to this inventory in order to define a subset of systems worthy of further detailing.

The following section described in detail first the selection criteria, and then the descriptors that it would be useful to collect for any system included within the case study application, an activity that could be compared to information domain modelling. For this section of the methodology, the author relied primarily on his professional experience in building and/or exploiting federated systems for various fisheries applications.

The methodology closed with a discussion of final validation steps, of which this document forms one part and the independent review the other.

Included as annexes were the interview templates, a spreadsheet rating the criticality of various system descriptors, and an XML schema to be used for system detailing.

3. Initial interviews

Beginning in mid-November, interviews (see Annex I – Interviews) were conducted with a group of fisheries experts with varying fish stock assessment backgrounds in order to create a seed set of systems that could be used as a basis for beginning the inventory. Ten interviews (including one auto-interview of the deliverable coordinator) were performed consisting of fisheries resource specialists working with traditional large data sets as well as specialists working with broader ecosystem-based models used more frequently in data-poor areas. Three of the ten interviews were with fisheries information system managers and/or developers.

4. Inventory creation and subset selection

As a result of the interviews, more than fifty different information systems were identified by fisheries experts. This group was then used to extrapolate a second group of similar sites, following the selection criteria outline in the methodology which led to a final list of 140 systems. For additional information on the system selection process please see Validate information systems.

The systems were entered into an MS-Access database (downloadable on the NeOn collaboration website: http://www.neon-project.org/ACollab/drafting/revisions.php?id=459) for the convenience of several team members in the United Kingdom who were familiar with the software. It was also felt that it would be easier to analyse the data if it were stored in relational tables as opposed to trapping it in a document such as this one.

The selection criteria found in the methodology were then applied to the overall set of systems, leading to a list of 33 systems deemed worthy of further detailing. For systems selected, the database contains a field describing the reasoning for system selection. For further information on systems selection please see Validate selected subset of information systems.

5. System detailing

For the selected subset, emails were sent to system managers informing them that their system was deemed useful to the FSDAS case study. It asked them for assistance in detailing the descriptors for their system that were identified by the methodology. As was to be expected, some replied, many did not, and responses are still trickling in. There were no negative replies.

It was decided that in addition to going forwards with systems that had replied positively, other systems would also be detailed using web-available information where use policies or practices indicated the data was for public consumption. This effectively reduced the number of detailed systems; however they can be added during the project lifecycle as more data or agreements come forth.

To facilitate systematic collection, an XML schema based on the methodology's system descriptors was created. It was also felt that XML files had strong advantages to either relational databases or document formats as they were both human and machine readable, could be easily circulated, were not dependent on any vendor-specific solution, and could be easily imported into databases and/or information systems.

Detailing began in early January and an XML file was elaborated for each system tackled. The XML detail files are not attached to this document as they run to over 350 pages. They can be found in the NeOn collaboration website at:

http://www.NeOn-project.org/ACollab/drafting/revisions.php?id=452

6. Inventory validation

This section covers the validation steps that were carried out during the inventory process, following the guidelines found in the methodology.

6.1 Validate information systems

Following the initial interviews the systems discovered were compared against the scope found in the methodology, *Section 1.2, Scope* and with the *Section 3, Candidate system selection criteria*. Additional systems were then selected in order to where possible expand the set of systems to meet the criteria. The table below shows the final breakdown:

System type	Covered
content management systems	TRUE
document repositories	TRUE
real-time data	TRUE
time-series statistics	TRUE
registries	TRUE
thesauri	TRUE
reference data systems	TRUE
websites	TRUE
forums/mailing lists	FALSE
rss feeds	TRUE

Assessment scale	Covered
National	TRUE
Regional	TRUE
Global	TRUE

Assessment facet	Covered
Biologic data	
species	TRUE
age	TRUE
mortality	TRUE
length	TRUE
Catch data	
landings	TRUE
biomass	TRUE
exports	TRUE
Effort data	
vessel type	FALSE
vessel size	FALSE
fishing gear type	FALSE
crew size	FALSE
trip duration	FALSE
trip distance	FALSE

Other considerations	Covered
Habitat	TRUE
Geo spatial	TRUE
Temporal	TRUE
Various approaches	TRUE
Various purposes	FALSE
Overall coverage:	**76%**

Figure 1. System scope coverage

The major area lacking in coverage was effort data, a dataset that is generally lacking in fisheries research and is a current area of interest. Although scientists are taking steps at the national level using vessel monitoring systems, this data is typically private as it touches on a very sensitive area for fishers.

The only other FALSE outcome was for *various purposes*, in which the available data for non-standard purposes for assessment were too few to warrant inclusion.

6.2 Validate selected subset of information systems

The list of 140 systems was weeded down to 33 systems by using the methodology *Section 3, Candidate system selection criteria* with an eye towards maintaining the broad scope of system types. The table below gives an overview of the reasoning behind 28 of the selected systems as the timeframe did not allow for the detailing of all 33 systems:

Name	RecommendCriteria	ActuallyDetailed	DetailedIn
AGROVOC Thesaurus	AGROVOC thesaurus is a great resource of keywords for fisheries and is already planned to be used for ontology learning experiments. It also contains many languages, important for testing NeOn functionality	1	AGROVOC.xml
Alaska Fisheries Science Centre - Auke Bay Laboratory	Subsite of NOAA. Has an interesting set of environmental variables keyed to species - unusual.	1	noaa_nmfs_site.xml
Aquatic Sciences and Fisheries Abstracts	ASFA is the resource for fisheries abstracts used widely by fisheries scientists and containing over one million records	1	ASFA.xml
Aquatic Species Fact Sheets	Hundreds of detailed factsheets together with catch statistics for species of major commercial importance	1	FIGIS.xml
ASFA Thesaurus	ASFA thesaurus is a great resource of keywords for fisheries and is already planned to be used for ontology learning experiments	1	ASFA_thesaurus.xml
ASFIS Taxonomic Authority List	ASFIS is the taxonomic backbone of the FIDI time-series datasets.	1	RTMS.xml
Australian Spatial Data Directory - Datasets for different fisheries in Australia	**Recommended but timeframe did not allow detailing**	0	
California Cooperative Oceanic Fisheries Investigations	**Recommended but timeframe did not allow detailing**	0	
Demand and Supply	**Recommended but timeframe did not allow detailing**	0	
FAO Document Repository	FAO corporate document repository containing thousands of official publications on fisheries.	1	EIMS.xml
FAO News and Events Systems (NEMS)	Active RSS feed with news on fisheries and events such as meetings of fisheries managers	1	NEMS.xml
FAOLEX	Registry of laws and regulations regarding fishing.	1	FAOLEX.xml
Fishbase	An encyclopedic site on fish species. The number one reference point for biological information about fish species	1	FishBase.xml
Fisheries Centre Working Papers Series	**Recommended but timeframe did not allow detailing**	0	
Fisheries Global Information System	XML-based queryable repository with fact sheets for a wide variety of fisheries domains.	1	FIGIS.xml
Fisheries Resources Monitoring System	Global authority for fish stock assessments made up of a large partnership of regional fishery bodies	1	FIRMS.xml
Fishery Statistics Programme	Global catch and production time-series statistics since 1950 used by many institutions for analysis	1	FIDISTAT.xml
Global Ocean Observing System	**Recommended but timeframe did not allow detailing**	0	
Globefish	Economic information such as market reports on commodity prices	1	GLOBEFISH.xml
Ministry of Fisheries	Nice mixture of stock catch data, commodities data, and management information for New Zealand waters.	1	New_Zealand_Min_Fish.xml

Name	RecommendCriteria	ActuallyDetailed	DetailedIn
National Marine Fisheries Service (Alaska)	Subsite of NOAA, good fishery resources data for NorthEast Pacific	1	noaa_nmfs_site.xml
National Oceanic and Atmospheric Administration (NOAA)	Main NOAA site. Large collection of data and subsites - wide array of data types.	1	noaa_nmfs_site.xml
Northeast Fisheries Science Centre	Subsite of NOAA, good fishery resources data for NorthWest Atlantic	1	noaa_nmfs_site.xml
oneFish	Community site with forums, news, management info, document links.	1	ONEFISH.xml
Poseidon	State of the art MIT site on ocean forecasting.	1	poseidon_site.xml
ReefBase	A good data component for ecosystem models. Unusual dataset.	1	ReefBase.xml
Reference Tables Management System	FIGIS reference terms management system	1	RTMS.xml
Resource Assessment and Conservation Engineering (Alaska Fisheries Science Centre)	Subsite of NOAA, good fishery resources data for Bering sea.	1	noaa_nmfs_site.xml
Resource Ecology and Fisheries Management Division of Alaska Fisheries Science Centre	Subsite of NOAA, good data on stock status for NorthEast Pacific.	1	noaa_nmfs_site.xml
Sea Around Us Project	Good example of a combination of GIS and time-series data.	1	theSeaAroundUs.xml
Southwest fisheries science centre (Environmental Research Division)	Good oceanographic datasets	1	noaa_nmfs_site.xml
Universal Biological Indexed and Organizer (uBio)	Looks to be the most technically advanced site dealing with taxonomic issues	1	uBio.xml
	Systems detailed=	**28 (85% coverage)**	

Figure 2. Systems selected for further detailing

6.3 Validate detailing for information system subset

Detailing was by far the most difficult task in the deliverable. Many system managers did not reply to our request for information, requiring additional work on the part of the team to determine if the data was publicly available, and then to determine if there was any public documentation of the API. In some cases it was necessary to rely on a scrape of public web forms. These problems caused the team to finish with a set of twenty seven detailed systems, some more complete than others.

The detailed systems were compared to the spreadsheet list of critical descriptors to determine to what extent these descriptors were covered. The following method was used:

1. Each detailed system was examined for the presence of descriptors, assigning a different rating depending on the criticality of the descriptor according to the methodology's *Appendix B*.

2. The lack of *any* critical category one descriptor generated a “problem”, as did the *complete* lack of a category two descriptor.
3. Warnings were generated for occasional missing descriptors at category 2 level and for completely missing descriptors at category 3 level.
4. All other cases resulted in complete validity.
5. An “overall validity” percentage was calculated by taking the overall descriptor set and dividing it against the set of valid descriptors.

Note that the weighting at file level did not weight the number of times a particular descriptor was present or not present, but rather looked for trends, such as “always present”, “sometimes present” and “never present”.

The table below summarises the results:

File	Validity %
UBIO	79
Sea around us	76
RTMS	84
Reefbase	74
Poseidon	74
OneFish	95
NOAA	89
New Zealand	68
NEMS	95
GlobeFish	82
Fishbase	82
FIRMS	97
FIGIS	87
FIDI	68
FAOLEX	89
EIMS	95
ASFA_thes	87
ASFA	92
AGROVOC	95

Average validity: 80%

Figure 3. Completeness of detailed systems

However, taking into account the fact that some files accounted for more than one system (e.g. NOAA, RTMS, and FIGIS), weighting these files by number of systems resulted in an **average validity of 86%**.

The complete results for each file can be found in a zipped file posted to the NeOn Collaboration website (http://www.neon-project.org/ACollab/drafting/revisions.php?id=460) as a series of HTML pages (XSL was used to build the results for each file).

7. Conclusions

The Inventory of Fisheries Resources and Systems should provide a sound basis for development of the WP7 case study. The application of a methodology and the relative completeness of the resulting data is a good sign that the inventory will meet needs.

Although a validation threshold was not included in the methodology, it seems clear from the abundance of detail data (350 pages of XML files) and the broad scope of detailed systems that there is more than enough upon which to base both the case study and any partner experiments.

As noted earlier, from the 28 systems detailed, 24 can be considered to meet validation requirements (above the 80% validation mean). For those systems not completely meeting validation, this was due to the impossibility of detailing them completely due to the non-interest of some system owners together with scant system documentation.

Further developments that might be considered to gain maximum benefit from the inventory would be to rework some of the early set of detailed systems in which parameter values when available were not always tokenized. A more complete tokenization of such values would facilitate the construction of a query component that could use the XML files to generate valid system queries based on the detailed interfaces. In addition, some detailed files that fell below the validation mean could still be extended time and/or need permitting.

8. Appendix A - Methodology

8.1 Overview

In order for the inventory of fisheries information systems to be useful, it should be backed by a coherent methodology that insures that the important aspects of every system are captured in a systematic way.

This document defines a process for the selection of candidate systems together with an expected scope of system types and content.

It goes on to define the parameters that should be considered for every inventoried system. Using these parameters it should be possible to identify and classify a set of NeOn-usable information systems relevant in making fisheries assessments that can be exploited in ontology-building and/or as data sources for the fisheries Assessment system case study.

Purpose

The purpose of the inventory is to capture a set of information systems that are likely to prove useful as part of the knowledge base for the WP7 case study “Fisheries Stock Depletion Assessment System”.

Considerations:

- Any inventory must logically be based on a subset of the data that is already used by Fisheries experts (We cannot expect to uncover data sources experts don't already know about. After all, they are the ones creating the body of data in the first place)

High-level data criteria:

- data needed by experts for assessments
- data needed by NeOn as proof-of-concept (use-case relevant)

Scope

8.1.1 Possible system types

- content management systems
- document repositories
- real-time data
  - markets
  - oceanographic data
    - temperature
    - salinity
- time-series statistics
  - capture
  - commodities
  - fleets
- registries
  - fleet/vessel
  - legal
  - fisheries management
- thesauri
- reference data systems
- websites
- forums/mailing lists
- rss feeds

8.1.2 Possible system domains

The criteria again should be based on data currently used by fisheries assessment experts when making assessments. Thus the initial interviews with experts are expected to define the limits of the possible domains. It is expected that the domains will include but not be limited to:

Oceanographic, ecological, other biological (e.g. taxonomy, physiology etc), hydrographic (e.g. CTD, current, ice, etc), meteorological (e.g. climate, wind, etc) and demographic data. See also 1.2.

8.2 Discovery, selection, inventory and validation methods

The method used to gather the inventory will be the following:

8.2.1 Discover candidate information systems

Fish stock assessments rely on three basic facets, and are created at three geographic scales. They may also have varying approaches and serve varying purposes. Each of these implies a different work flow, a different basic set of data and information system, and perhaps a different professional background. Interviews with fish stock assessment experts will be used to determine the set of systems which may be inventoried for the FSDAS case study. As such, it is important to interview experts working with various facets and at various levels in order to be sure of having a representative sample of information systems.

8.2.1.1 Procedure

- Interview (live, phone or by email) fish stock/resource assessment experts. These experts will be selected based on the following criteria:
  - experts at FAO within the Marine Resources Service (FIRM)
  - experts at regional fisheries bodies that participate in the Fisheries Resource Monitoring System (FIRMS)
  - experts at national level that work with facets, approaches and/or purposes that are otherwise under-represented in the inventory (see section Candidate system selection criteria)

  (See Appendix A for a sample interview form)
  - Using the systems generated by the interviews as a basis, select similar systems with an eye towards meeting as much as possible the scope of systems set out in the Candidate system selection criteria)

8.2.2 Validate information systems

Once a set of systems has been identified for an initial inventory, it will be necessary to validate its scope based on the criteria set out in the Candidate system selection criteria.

8.2.2.1 Procedure

For each system, compare it against the scope found in the **Candidate system selection criteria**

8.2.3 Select subset of candidate information systems for further detailing

Based on the criteria set out in Candidate system selection criteria a subset of the inventory will be selected for further detailing based on their perceived usefulness to the case study and to NeOn project partners.

8.2.3.1 Procedure

- For each system in the overall inventory, apply criteria described in Candidate system selection criteria to identify the most relevant systems, listing for each selection the motivation.

8.2.4 Validate selected subset of information systems

Once a subset of systems has been identified it will be necessary to validate the list based on the criteria set out in the Candidate system selection criteria.

8.2.4.1 Procedure

- For each selected system, compare the listed motivations against the Candidate system selection criteria.

8.2.5 Detail subset of candidate information systems

Once a set of candidate systems has been defined we need to describe them such that we know what kind of data and metadata they hold, how the data can be accessed, access rules and rights, and whether the system owners will either let us use their data or are willing to work with us in some way. The section, Information system description, groups and defines aspects of systems that we are interested in knowing something about for the inventory.

Using the candidate set of systems we will contact system managers and/or system developers to gather this descriptive data. In many cases, we may be unable to gather the entire set. This is okay. It is not expected that we will be able to; the aspects presented here are more of a wish list in order to make sure we aren't leaving some aspect uncovered. For relative importance of descriptors see APPENDIX B – Relative importance of system descriptors

8.2.5.1 Procedure

- Once candidate systems have been selected these systems will be inventoried by gathering descriptive data as outlined in section Information system description. System administrators and/or developers will be interviewed (live, phone or by email).
- Data will be organised and stored in either a spreadsheet, XML file or database table to ease future use.

8.2.6 Validate detailing for information system subset

Once detailing is complete it will be necessary to determine whether the detailed inventory covers the descriptors identified in Information system description.

8.2.6.1 Procedure

- Once selected systems have been detailed, compare detailing for each system against the descriptive data outlined in section Information system description.

8.3 Candidate system selection criteria

The following sections offer a detailed set of facets, scales, approaches and purposes that should be considered both when selecting interview candidates and when selecting candidate systems.

Assessment facets

- Biologic data
  - species
  - age
  - mortality
  - length
- Catch data
  - landings
  - biomass
  - exports
- Effort data
  - vessel type
  - vessel size
  - fishing gear type
  - crew size
  - trip duration
  - trip distance

Assessment scale

- National
  - These may range from nothing more than raw sampling data forwarded to regional level, to large-scale sophisticated assessment programs.
  - Multiple countries may report together on shared stocks, e.g. Italy and Croatia for Adriatic stocks.
- Regional
  - Typically prepared by regional fisheries bodies, these reports bring together both raw data and prepared assessments from national level.
- Global
  - Typically prepared from national and regional reports, they are overviews that cover large water areas.

Other considerations

There are other considerations that may impact the way in which assessments are performed and the kinds of information systems and/or source data used.

- Habitat (inland, coastal, deep sea)
- Geo spatial, temporal (inland lentic fisheries –vs.- marine)

Approaches

Whereas section Assessment facets presents a traditional assessment approach there are other more recent approaches that imply different data and assessment methods that deserve examination such as the “ecosystem approach” that assesses a fish stock in relation to its environment, e.g. other species, food, predators, habitat, etc. There may be other approaches that emerge as initial interviews are conducted.

Purposes

Assessments are carried out in order to indicate the current state of a fish stock, both status and trends. However, there are various purposes for which an assessment is carried out. While

assessments are typically carried out in order to help managers in making management recommendations or regulations designed to maintain stocks at sustainable levels, there are also assessments carried out for other purposes. For example, the CITES convention on endangered species requires that nations exporting species on the list have carried out an assessment indicating that the current landings are at a sustainable level.

8.4 Information system description

This section describes all the aspects we hope to cover for the systems we finally inventory. Many are really part of a wish list, while others are crucial. See Appendix B for a matrix indicating the importance level of each descriptor.

8.4.1 General aspects

Here we work at a level above the information system itself. How does this system fit into the overall inventory? Why are we interested in it?

8.4.1.1 Selection criteria

- Why was the system selected for inclusion?
- Which kind of experts uses the system in their work?
- Who are the typical users of the system and how do they use the system?
- In what broad ways does it satisfy the needs of the case study?
- In what broad ways does it satisfy the needs of NeOn project partners?
- Is the system authoritative?
- Is the system encyclopedic?

8.4.1.2 System relations criteria

- How does this system relate to other systems?
- Are there applications that use this system together with other systems to generate a synthetic output?
- Is there overlap between this system and other systems in the inventory?

8.4.2 Data aspects

Here we consider all the questions related to the data itself, outside of questions of how it's collected, stored, disseminated, etc. This is really separate from any system or application that stores, updates or uses the data.

8.4.2.1 Data

Here we consider the raw data itself, outside of other information that may describe it. Primarily we want to know what dimension it covers, together with how it is expressed.

Coverage

- What is the coverage of the data subject area?
  - Spatial

- Temporal
- Species
- Effort
- Language
- Overall size in storage terms

Resource Type

- Statistical time-series
- Document-like data
  - PDF, HTML, Word, DocBook, XML
- Structured markup
  - XML, SGML

8.4.2.2 Metadata

What kind of metadata such as classification systems are used to classify the data, these could possibly be identifiers attached to the data or column headings in a database table, or XML elements or attributes, or META tags in html documents.

- What kind of classification system?
- Maintained by whom?
- How updated?
- How widely used?
- How expressed?
  - Metadata schemas
  - Other?

8.4.2.3 Quality assurance data

Quality assurance data can also be thought of as a type of metadata: provenance, validity and collection method.

Collection methodology

- Where does data come from?
- How is data collected?
- How frequently?
- For how long?
- By whom?
- According to what protocols?
- How verified? Random re-sampling?

Post-collection

- How processed and handled?
- How is it input into an information system?
- How is it verified? Double entry?

Authority

- Who stands behind the data validity?
- What is their authority?

8.4.2.4 Data structure

Here we go to a higher level of abstraction and consider how the data is organised. In software development terms we are going to a design stage. What is the basic organisational paradigm?

Hierarchical schema-based

- XML and SGML type of data structures.

Relational tables

- Likely structure for time-series data. Some systems may hold document-like objects.

Ontology-based

- Knowledge-base with instances attached to ontological concepts. Generally unlikely.

Object-oriented

- Persistent data store based on OO concepts. Generally unlikely. But in a system that uses remote objects via CORBA or RMI this may be a viable structure to consider.

8.4.3 Legal aspects

Questions surrounding the use of data. Who can access it? Is it public? Is it pay-per-use? What are the restrictions on its use?

8.4.3.1 Access

- What are the access rights?

8.4.3.2 Copyright

- What are the rules if any for attributing ownership?
- What are the rules if any for citation?

8.4.4 Technical aspects

Here we consider how the information system is put together and accessed. In software design terms this would be akin to the stage of actual implementation.

8.4.4.1 How can the data be accessed or exposed?

What are the protocols by which data can be accessed? What is the paradigm? Message passing via web services using TCP/IP or some custom protocol? By using distributed object technology such as CORBA or RMI? Queries using HTTP? How is the data returned? In some native proprietary format such as PDF or Word? As XML? CSV?

8.4.4.2 What are the APIs?

Here we specify more closely the aspects of data access. Is there an interface that is known, documented and stable? When it changes, is support maintained for previous methods? Are there deprecated methods that are phased out over several releases, etc.?

- Are they published?
- Are they stable?
- How is change handled?
- Are they extensible?

8.4.4.3 How efficient is data retrieval?

Here we consider performance issues and issues of change over time. How often do we need to poll a system. Is there any kind of event-based service we can subscribe to as an observer that will inform us of change? How hard can we load a system with data requests? Are data structures and models fairly stable over time? When they change how is this fact broadcast?

- What kind of load can the system handle?
- How often does the data change?

8.4.5 Administrative aspects

These questions relate to the environment in which the system exists. Typically these would be requirements level questions when building a system. We need to know the kind of world the system lives in. Often these questions relate to stakeholders that have some interest in or authority over the system but are not developers, editors or users.

8.4.5.1 Who has authority for the system?

Who needs to be consulted when change or collaboration or new developments are proposed?

8.4.5.2 Who can change the data?

Which person(s) or group(s) have the authority to change the data and for what reasons? Is there a change strategy in place at management level?

8.4.5.3 Who can modify the system?

Who is allowed to actually modify the functionality or system/data model and for what reasons?

Practical aspects

Finally we consider questions which relate to how outside groups can interact with the system, in our case NeOn. We need to know in a simple kind of way if we can have a working relationship, and if so to what degree.

- Are they interested in working with the NeOn case study?
- Can they grant access to the data?
  - Any restrictions or conditions?
- Can they modify or extend the API if needed?
- Can they modify the data structure if needed?
- Do they have resources to expend for interaction with NeOn?
- Do they have resources to expend were there a need for API modification/extension?

8.5 Final validation

Although sequential validation steps have been described in the method section, final validation should consist of the following steps:

- A results document that concretely describes the process followed in the execution of the inventory exercise.
- An independent review of the deliverables to establish that they meet the standards set forth in this document, and more generally for a professional inventory exercise of this nature.

Time permitting, it would be instructive to use several detailed systems in an actual implementation of a knowledge base in order to verify that the descriptors were indeed sufficient for real use, but the brevity of the deliverable does not permit this step.

Annex I – Interviews

See next pages.

Note: The interviews are transcriptions of raw field notes and data and as such may contain spelling, grammatical and/or other errors. The templates used during interviewing were modified over time such that there may be slight differences in template design from one interview to another.

Inventory Interview Form

Name: Fabio Carocci

Email: Fabio.Carocci@fao.org

Webpage: Tuna Atlas

Areas of expertise: GIS, time-series

Regions/nations: Global, Japan, USA (NMFS)

Regional fisheries bodies: IOTC - Indian Ocean Tuna Commission

SPC (the Secretariat of the Pacific Community) CCSBT (Commission for the Conservation of Southern Bluefin Tuna) IATTC (the Inter-American Tropical Tuna Commission) ICCAT (International Commission for the Conservation of Atlantic Tunas)

Data facets worked with:

- Biologic data (species, age, mortality, length)
- Catch data
    - landings – how much, which stock, who caught it
    - capture - 5 degree square data on how much by species group
- Effort data (vessel type, vessel size, fishing gear type)
    - for captures, they know how much of a species group by gear, when and where

Assessment scale performed:

N/A

Approaches used:

N/A

Assessment purpose:

N/A

reports/publications/data regularly PRODUCED:

name:	Tuna atlas
type:	maps
publishing frequency:	annual
url if any:	http://www.fao.org/figis/servlet/TabSpecies?tb_ds=TunaAtlas&tb_mode=MAP&tb_act=SELECT&tb_grp=SPECIES

name:	Global tuna nominal catches
type:	time-series
publishing frequency:	annual
url if any:	http://www.fao.org/figis/servlet/TabStock?tb_ds=TunaLandings&tb_mode=TABLE&tb_act=SELECT&tb_grp=STOCK

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	

url if any:	

reports/publications/data regularly CONSUMED when producing:

name:	GeoNetwork
type:	GIS
url if any:	http://www.fao.org/geonetwork/srv/en/main.home
contact name:	MARTUCCI, Antonio (SDRN)
contact email:	Antonio.MARTUCCI@fao.org
contact phone:	
contact organisation:	FAO - SDRN

name:	Indian Ocean Tuna
type:	spreadsheets
url if any:	http://www.iotc.org/English/data/databases.php
contact name:	
contact email:	secretariat@iotc.org.
contact phone:	
contact organisation:	IOTC - Indian Ocean Tuna Commission

name:	Pacific Tuna
type:	database tables
url if any:	http://www.spc.int/oceanfish/html/statistics/index.htm#Public
contact name:	
contact email:	mailto:ofp@spc.int

contact phone:	
contact organisation:	SPC (the Secretariat of the Pacific Community)

name:	Southern Bluefin Tuna
type:	spreadsheets
url if any:	http://www.ccsbt.org/docs/data.html
contact name:	Robert Kennedy
contact email:	rkennedy@ccsbt.org
contact phone:	
contact organisation:	CCSBT (Commission for the Conservation of Southern Bluefin Tuna)

name:	Tropical Tuna
type:	PDF
url if any:	http://www.iattc.org/DataENG.htm
contact name:	
contact email:	
contact phone:	
contact organisation:	IATTC (the Inter-American Tropical Tuna Commission)

name:	Atlantic Tuna
type:	Spreadsheet, database
url if any:	http://www.iccat.es/accesingdb.HTM
contact name:	

contact email:	
contact phone:	
contact organisation:	ICCAT (International Commission for the Conservation of Atlantic Tunas)

Notes:

Carocci works primarily with GIS systems and their application to statistical data. He described the set of maps used within FAO as the following:

UN level maps: National Areas

FAO: Land use, rivers, lakes

FI/FAO: base maps of coastline and bathymetry at various scales

FI: Fishing areas, Regional body competence areas

He discussed GeoNetwork, an in-house system that holds GIS layers for FAO.

He also noted that the English CEFAS had excellent GIS and other hydrologic / oceanographic data.

Carocci mentioned a very interesting project using a combination of VMS (vessel monitoring systems), landings and logbook data to monitor crustacean fisheries: http://w3.ualg.pt/~madias/geocrust/index.html. It represents a very cutting-edge approach to using GIS and monitoring systems together with traditional landings data.

Lastly he added that apparently there were good data at the National French site: IFREMER though he had not personally made use of it.

Inventory Interview Form

Name: HALL, Stephen John

Email: s.hall@cgiar.org

Webpage: http://www.worldfishcenter.org/cms/default.aspx

Areas of expertise: Ecosystem effects of fishing activities - Fishery resources monitoring

Regions/nations: Global and Regional (Asia)

Regional fisheries bodies:

Data facets worked with:

- Trawl surveys
- Fisheries ecology data
- Effects of fishing on the environment

Assessment scale performed:

- National YES for some countries
- Shared
- Regional YES Asia
- Global YES

reports/publications/data regularly PRODUCED:

name:	TrawlBase
type:	Repository of trawl data from Asian governments – between 20,000 and 40,000 trawl surveys
publishing frequency:	
url if any:	http://www.worldfishcenter.org/trawl/first/first/first.asp

name:	ReefBase
type:	Document repository
publishing frequency:	
url if any:	http://www.reefbase.org/reeffisheries/default.aspx

name:	Demand and Supply
type:	Document repository
publishing frequency:	
url if any:	http://www.worldfishcenter.org/demandsupply/output.htm

name:	FishBase
type:	Encyclopedic
publishing frequency:	
url if any:	

name:	oneFish
type:	Content management system

url if any:	

reports/publications/data regularly CONSUMED when producing:

name:	A very wide range of documents and databases – too many to number.
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

Notes:

Sources drawn on: extremely diffuse – mostly not in repositories or on web sites. The most important sources are synthesis documents. Raw data needs expert interpretation.

Listing all these sources is a non-trivial activity if it is to be meaningful. Longer time scales for collecting this sort of information would be required to make the exercise more meaningful.

The ecosystem approach to fisheries management is still in the early stages of assessing rules and values. Most work to date has been on Continental Shelf fisheries (but not “near-shore”) and large estuaries. The deep water work has mainly been on single species e.g. tuna.

WorldFish will assist the NeOn project with its publicly available information systems – ReefBase and “Demand and Supply” if this adds value and helps fulfil WorldFish’s objectives. There are definite limits to the amount of time they can give (hours not days). All requests should go through Stephen and not direct to systems people. Unfortunately, TrawlBase is a “No Go” area – each individual government department would have to be approached to get access to their data otherwise WorldFish would breach its agreed access protocols for this data.

Inventory Interview Form

Name: Yves Jaques

Email: yves.jaques@fao.org

Webpage: http://www.fao.org/fi/figis

Areas of expertise: software developer

Regions/nations: n/a

Regional fisheries bodies: n/a

Data facets worked with:

- Biologic data (species, age, mortality, length)
- Catch data (landings, biomass, and exports)
- Effort data (vessel type, vessel size, fishing gear type, crew size, trip duration, trip distance)

Assessment scale performed:

n/a

Approaches used:

n/a

Assessment purpose:

n/a

reports/publications/data regularly PRODUCED:

name:	HSVAR
type:	vessel authorisation records
publishing frequency:	annual
url if any:	http://www.fao.org/figis/hsvar/index.jsp

name:	FIRMS stock assessment fact sheets
type:	stock assessment fact sheets
publishing frequency:	annual
url if any:	http://firms.fao.org/

name:	Vesseltype fact sheets
type:	fact sheets
publishing frequency:	n/a
url if any:	http://www.fao.org/figis/servlet/static?dom=root&xml=tech/vessels_search.xml

name:	Geartype fact sheets
type:	fact sheets
publishing frequency:	n/a
url if any:	http://www.fao.org/figis/servlet/static?dom=root&xml=tech/gears_search.xml

name:	Fishing techniques
type:	
publishing frequency:	

url if any:	http://www.fao.org/figis/servlet/static?dom=root&xml=tech/fishtech_search.xml

reports/publications/data regularly CONSUMED when producing:

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

name:	
type:	
url if any:	
contact name:	
contact email:	

contact phone:	
contact organisation:	

Notes:

Interesting information systems at FAO that have not otherwise been covered in other interviews or in the above tables:

Fleet statistics:

to 1995:

http://www.fao.org/figis/servlet/TabLandArea?tb_ds=Fleet&tb_mode=TABLE&tb_act=SELECT&tb_grp=COUNTRY

1996-1998:

decked:

http://www.fao.org/figis/servlet/TabLandArea?tb_ds=Decked&tb_mode=TABLE&tb_act=SELECT&tb_grp=COUNTRY

undecked:

http://www.fao.org/figis/servlet/TabLandArea?tb_ds=Undecked&tb_mode=TABLE&tb_act=SELECT&tb_grp=COUNTRY

Fishery country profiles:

http://www.fao.org/figis/servlet/static?dom=root&xml=geography/fishery_CP.xml

Fisheries commodities:

http://www.fao.org/figis/servlet/TabLandArea?tb_ds=Trade&tb_mode=TABLE&tb_act=SELECT&tb_grp=COUNTRY

Species distribution maps:

http://www.fao.org/figis/servlet/FiRefServlet?ds=subservlet&session=Species&xsl=webapps/figis/subservlet/species/selectone.xsl&url=figis/kimsmaps/species[3Fquery[3DsNames[2CcNames[26outformat[3Dxml

EIMS search interface:

(FAO document repository)

http://www.fao.org/figis/servlet/static?dom=root&xml=refer/pubsearch.xml

NEWS and EVENTS search interface:

http://www.fao.org/figis/servlet/static?dom=root&xml=refer/nemssearch.xml

Fisheries RSS Feed:

http://www.fao.org/nems/rss/rss_nems_results.asp?owner=1&official=1

FAOLex:

(fisheries legislation)

http://faolex.fao.org/faolex/

ONEFISH:

http://www.onefish.org/global/index.jsp

UN Atlas of the Oceans:

http://www.oceansatlas.org/index.jsp

Globefish:

http://www.globefish.org/

Inventory Interview Form

name: Lamboeuf, Michel (FIRM); Lleonart, Jordi (FIRM)

email: Michel.Lamboeuf@fao.org; Jordi.Lleonart@fao.org

webpage: SIDP programme, Copemed, Adriamed, GFCM, CECAF, ICCAT

areas of expertise: marine biology, stock assessment

region: Mediterranean

Regional Fisheries Bodies: GFCM (Copemed) and CECAF

Standards: MEDFISH

Notes:

There are three main components to fish stock assessment:

- Biological – length, age, maturity
- Catch – landings by weight or piece
- Effort – number of vessels, kinds of gear, length of trips

Assessments are followed by management recommendations that perform cost/benefit analyses in possible approaches to reducing effort, creating marine protected areas, etc. This is not a part of assessment.

For the Biological component, data collection techniques are sampling, either onboard or at ports (length, age, catch composition and quantity).

For the Catch component, techniques are trawl and acoustic sampling for biomass/unit of area and catch composition, sampling in ports and from commercial sales. Artisinal fishing is not generally reported. Catches are generally extrapolated from sampling. In some cases they may be exhaustive depending on how formal the sales markets are.

Jordi Lleonart works with GFCM and described the makeup of the GFCM which manages mediterranean fisheries:

1. Countries create reports at national level for stocks in their national waters. (Biologists and Statisticians)
2. In some cases countries take their national reports and harmonise them to prepare reports on shared stocks, such as Italy and Croatia for the Adriatic, often via projects such as Adriamed and Copemed.
3. These reports known as “working group reports”. They are public. Their sources as well if they are listed are public.
4. These reports are given to the GFCM Stock Assessment Committees (of which working groups may form a part).
5. The committees may endorse, amend or reject working group reports.
6. The committees prepare scientific reports based on the working group reports which then carry the GFCM stamp.
7. These reports are then the basis for management decisions made by GFCM which carry legal authority.

Michel Lamboeuf who is more acquainted with West African fisheries described a different process, in which a meeting is held by a regional body, for example CECAF, who reports on Eastern Central Atlantic fisheries. They invite who they think will actually be helpful from each country. Typically several outside experts are there as well to help harmonise data that is brought.

Both experts noted that some countries are reluctant to give their data, perhaps because they know that it is low quality.

They noted the difficulties of assessing high-seas stocks such as large pelagics as they are in international waters.

Noted that sometimes two RFB’s may monitor the same stock, such as Tuna in the mediterranean, which is monitored by both GFCM and ICCAT, though in cases such as these one group is designated the lead (ICCAT in this case)

Noted that Effort data is a can of worms. Groups have been trying to decide for years how to standardise such data (known as “operational units”).

reports/publications/data regularly produced:

name	type	first pub	pub freq	url (if any)
SIDP Fact Sheets	species	1970's	annual	http://www.fao.org/figis/servlet/static?dom=root&xml=species/species_search.xml

Description: xml-based fact sheets containing biological, habitat, capture statistics and distribution for 548 commerically important species.

information systems/data sources used:

name	type	first_pub	pub_freq	contact	url (if any)
FIDI statistics	time series	1950's	Annual	luca.garibaldi@fao.org	http://www.fao.org/figis/servlet/static?dom=root&xml=tseries/index.xml
KIMS mapping system	GIS	2002	Dynamic	fabio.carocci@fao.org	http://www.fao.org/figis/servlet/static?dom=root&xml=maps/index.xml
RTMS	ref data	2000	Annual	francesco.calderini@fao.org	http://www.fao.org/figis/servlet/RefServlet

Inventory Interview Form

Name: Piero Mannini

Email: Piero.Mannini@fao.org

Webpage: N/A

Areas of expertise: assessment

Regions/nations: Mediterranean

Regional fisheries bodies: Adriamed, GFCM

Data facets worked with:

- Biologic data (species, age, mortality, length)
- Catch data (landings, biomass, and exports)
- Effort data (vessel type, vessel size, fishing gear type, crew size, trip duration, trip distance)

Assessment scale performed:

- National
- Shared
- Regional

Approaches used:

Traffic light method, polyvalent. Seek to implement fish management based on common standard of **effort**

“responible fisheries approach”

Assessment purpose:

Adriatic sea management

reports/publications/data regularly PRODUCED:

name:	medsudmed
type:	regional ecosystem project website
publishing frequency:	
url if any:	http://www.faomedsudmed.org/

name:	adriamed
type:	project website
publishing frequency:	
url if any:	http://www.faoadriamed.org/

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	

url if any:	

reports/publications/data regularly CONSUMED when producing:

name:	world conservation union
type:	website
url if any:	http://www.iucn.org/places/medoffice/en/en_fishery.html
contact name:	
contact email:	francois.simard@iucn.org
contact phone:	
contact organisation:	

name:	copemed
type:	website
url if any:	http://www.faocopemed.org/en/index.htm
contact name:	
contact email:	csirke@fao.org
contact phone:	
contact organisation:	

name:	Fisheries Centre
type:	university
url if any:	http://www.fisheries.ubc.ca/
contact name:	
contact email:	office@fisheries.ubc.ca

contact phone:	
contact organisation:	UBC

Notes:

Mannini has been behind the adoption of a standard way of measuring effort to allow for common management of the mediterranean/adriatic. The approach is known as “medfish”(http://www.fao.org/DOCREP/005/Y2793E/y2793e04.htm) and is still a proposal.

Medfish:

“standardized vessel, catch and effort data”

The concept is that of defining “operational units” based on “segments”

A segment is a vesseltype/geartype/area

The idea is also to know when they operate, what they catch, the makeup of the crew, length of trips, age of fleet.

The medfish information system is in development and will be built via GFCM RFB.

They are working so far on having vessel data on all boats above 50 metres authorised to fish in the mediterranean.

This will be put together with FAO/FIDI statistics on production viewed as a ten year trend compilation. Should be done in the next year or so.

Traffic light method

As opposed to traditional assessment methods, the traffic light method attempts to weight many more indicators (http://www.fao.org/docrep/006/y5029e/y5029e0d.htm) . Mannini noted that such an approach is conducive to our case study.

This means using a set of “multi-disciplinary indicators” and not just the “application of models”. Particularly looks at the conjunciton between market indicators and fishing effort. That is “market/people/stock/fleet”.

Suggested resources:

Mannini suggested trying to do a screening of projects for the last twenty years by scanning websites of various fisheries ministries, regional fishery bodies, univeristies etc. He said that much data and analysis is trapped in such obscure sources, many for which the project was long ago closed.

He also suggested searching out “fisheries economic analysis” (but an attempt on Google was not too promising).

Lastly, he suggested seeing possible reports created by important NGO’s such as IUCN and WWF.

Inventory Interview Form

Name: Taconet Marc

Email: marc.taconet@fao.org

Webpage: http://www.fao.org/figis/servlet/static?dom=root&xml=index.xml

Areas of expertise: Fishery resources monitoring – Fishery Information Systems development

Regions/nations: Global

Regional fisheries bodies:

Data facets worked with:

- Biologic data (species, age, mortality, length)
- Catch data (landings, biomass, and exports)
- Effort data (vessel type, vessel size, fishing gear type, crew size, trip duration, trip distance)

Assessment scale performed:

- National YES for some countries
- Shared YES in some cases
- Regional YES
- Global YES

reports/publications/data regularly PRODUCED:

name:	Fishery resources monitoring system (FIRMS)
type:	Monitoring status of fishery resources from global to national level (to be extended with Fisheries and their management) FIRMS interest with IODE/GeBich members evolves from increasing focus on the Ecosystem Approach to Fisheries: FIRMS needs to share standards on environmental and biodiversity subject matters; and also needs to facilitate access to relevant environmental information. taxonomic classifications, ecosystems descriptors (eg environment / habitat classifications, predator / prey relationships, measures and qualifiers of biodiversity, or primary productivity).
publishing frequency:	mainly yearly
url if any:	http://firms.fao.org/figis/website/FIRMSRetrieveAction.do?xml=FIRMS_org.xml&dom=org&xp_nav=1

name:	Fishcode-STF website
type:	Inventory of resources and fisheries (coming soon in 2007)
publishing frequency:	once off
url if any:	http://www.fao.org/figis/servlet/static?xml=STF_proj.xml&dom=org

name:	FIRMS members websites
type:	Assessments of Regional resources/stocks, Management measures for related fisheries
publishing frequency:	
url if any:	access to FIRMS members websites from http://firms.fao.org/figis/website/FIRMSRetrieveAction.do?dom=org&xml=FIRMS_org.xml&xp_nav=2,3

name:	Ocean Biogeographic Information System (OBIS)
type:	OBIS is the information system of the Census of Marine Life (CoML) programme. CoML aims over a 10 years programme at understanding past, current and future patterns of marine life. OBIS, a distributed web-based one, serves this purpose by collating from a great diversity of sources (museum collections, existing systems like Fishbase, scientific cruises at sea, ...) georeferenced records of species occurrence and disseminating this knowledge through dynamically generated maps or downloadable files. The challenge is to enumerate all taxa living

	in the oceans, the current estimated figure being that of 160.000 species, and to better understand marine biodiversity.
publishing frequency:	
url if any:	http://www.iobis.org/OBISPortal/

name:	USA: NMFS Species Information System (SIS)
type:	the NMFS Species Information System (SIS) is a FIGIS sister system aimed at producing the USA status of stocks report (*Our Living Oceans*) and the *Status of the U.S. Fisheries* report to Congress.
publishing frequency:	
url if any:	

name:	Canada: DFO statistical and quota monitoring system and the integrated fisheries management administrative support (FIMP) system.
type:	Existing CSAS system features are impressive. Trip data from all Canadian vessels are available in near-real time at DFO headquarters. The system is essentially designed to provide real-time administrative support to management. CSAS's static web site implements a systematic and transparent reporting for the various disciplinary approaches to fisheries management – biological, ecosystem, and socio-economic. It appears that the Canadian management framework is clearly in place, but that it has not been formalised through a system like FIGIS. For example, there are no systematic inventories managed in a database way, allowing interactions between all the participants to knowledge elaboration and decision making.
publishing frequency:	
url if any:	

name:	Australia: BRS and ABARE
type:	BRS maintains a detailed Fishery status report for 13 commonwealth resources, accessible individually on the web. Out of the 6 Australian states, 3 (Western Australia, ? and to a certain extent new South Wales) publish in paper format State fisheries status reports. The next BRS 67 commonwealth species.

publishing frequency:	
url if any:	

name:	Regional fishery bodies websites
type:	see list of RFB members of FIRMS at http://firms.fao.org/figis/website/FIRMSRetrieveAction.do?dom=org&xml=FIRMS_org.xml&xp_nav=2,3
publishing frequency:	
url if any:	

name:	SEAFDEC website
type:	o The SEAFDEC web Digitized Atlas (Regional Fishery Information System) is a response to the need to provide a single entry point to SEAFDEC databases. The Training Department/Research Division, which has a computer system capacity is responsible for its development. The Atlas provides access to data structured during different Projects: Sea base (collection of oceanographic data collected through collaborative cruise surveys program); sea turtle projects (distribution of turtles, nesting locations, fishing grounds, tagging results); Inland Capture (environmental conditions, water body list, relevant regulations, and definitions, by country); Tuna research project (study of environmental factors explaining the abundance of Tuna, through cruise survey data); coastal management project; and indeed Regional Fishery Statistics.
publishing frequency:	
url if any:	

reports/publications/data regularly CONSUMED when producing:

name:	
type:	

url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

Notes: Extracted from my Back to Office Report to Gebich3 meeting, Oostende

The NeOn project could provide an opportunity to test, through integration in the NeOn ontology of fisheries, taxonomic, and oceanographic schemas, related classifications and APIs, how well NeOn tools could enable underlying systems to interoperate. It was noticed that NeOn could be a possible tool to implement the IODE/Gebich action entitled “Investigate tools to create distributed Metadata catalog” lead by Edward Van Den Berg and Roy Lawry, and that should the concept prove able to respond to the needs, IODE could consider if/how funding provisioned for this action could be contributing to the broader effort on Ontology development.

The relevant schemas which should in particular be considered:

Metadata repositories and schemas:

- MEDI: the Marine Environmental Data Inventory of the IODE metadata system

- ISO 19115 with the extended Marine community profile

Taxonomic repositories and schema:

- TSC: Taxonomic schema (developed under the TDWG group)

- ITIS, the Integrated Taxonomic Integrated System that the GeBich group wishes to become the international authoritative system for Marine species

- APHIA, a repository of taxonomic lists (including UNESCO Register of Marine Organisms (URMO), ASFIS, ERMS, Global Databases on Species (GDS)) which will eventually provide the foundation for the envisaged World Register of Marine Species (WoRMS), seen as the future subset of ITIS

Bio/ecological systems and associated schemas backed by protocols:

- BIOCASE schema and protocol

- DIGIR schema and protocol (used by OBIS)

- TAPIR schema and protocol (TAPIR would eventually substitute BIOCASE)

- SOAP

- ICES website: fisheries working group reports, other science document repositories about ecological-environment-oceanographic-ecosystems information

- Ontology of Marine oceanography parameters (Roy Lowry)

From the work of TDWG (www.tdwg.org) : Taxonomic Database Working Group, now renamed **TDWG Biodiversity Information Standards**

has established a TDWG ontology looking at first class biodiversity objects and their relations

TAPIR schema: merges DiGir and Biocase protocols. TAPIR should replace the two other protocols. Integrating these schemas in NeOn and underlying protocols would be great since it would enable interoperating with systems like OBIS (which uses the DiGir protocol)

the **Marine XML group**, led by Roy Lawry, remains responsible for all Marine vocabulary. This group develops the controlled Marine vocabulary through the SeaVox mailing list in which anybody interested can subscribe.

World Register of Marine Species (WoRMS)

WoRMS is to become the universal register of Marine species. It will be built from an assemblage of sources to ensure its comprehensiveness, and its records will be cleared by assigning ITIS codes. 120.000 taxonomic records have already been gathered in the APHIA repository available at www.vliz.be/vmdcdata/aphia

The idea with WoRMS is to create a sort of subsidiary body of ITIS (through a protocol arrangement) for marine species in order to remove the recorded bottleneck of clearing records at ITIS level. The WoRMS will be developed through a specific project, and the work will consist in validating WoRMS records through assigning ITIS Ids for those records mapped with existing ones in ITIS, all the other records requiring a specific validation. Problems are essentially anticipated in Crustaceans and Molluscs.

UBIO is a set of intelligent tools to help identify the exact taxa for a given record. For example when applied to APHIA it helps to overcome current typing mistakes in species names.

The proposal is to consider how NeOn can play a role in constructing WoRMS. A preliminary step would be to map ASFIS list of species with the APHIA list, thus assigning ASFIS codes and FAO names to APHIA.

In order to help the NeOn project in decision making, Edward Van den Berg will look at the **TCS** schema (which describes the Taxonomic data model) and verify is APHIA can be extracted to TCS. He will also look how many records from ASFIS are there and how they match with APHIA records list and OBIS list, and finally will report those results to NeOn project.

Systems or data bases considered for constructing WoRMS:

ASFIS list of species for fishery statistics purposes

ITIS: Integrated Taxonomic Information System

URMO: UNESCO Register of Marine Organisms

ERMS: European Register of Marine Species

OBIS: Ocean Biogeographic Information System

GSD: Global Species databases, eg Fishbase

Catalog of life: fed from Species 2000;

= GSD + IT IS in 2006,

= GSD + IT IS + ERMS in 2007

North Atlantic Register of Marine Species (NARMS)

This is a Canadian data base which Bob Branton (now fully working on taxonomic classifications) strongly promotes. Available at www.vliz.be/vmdcdata/narms/browther.php.

It should be noted that Bob Branton has already mapped ITIS and ASFIS codes for Canadian species recorded during Canadian scientific surveys.

Inventory Interview Form

Name: Merete Tandstad
Email: Tandstad.Merete@FAO.ORG
Webpage: n/a
Areas of expertise: stock assessments
Regions/nations: West Africa, NW Africa
Regional fisheries bodies: CECAF

Data facets worked with:

Primarily catch data from nations.

Observer data – national scientists may go on any fleets boats that are in their national waters.

Some regional biologic data from independent surveys (acoustic and/or bottom trawl).

Some effort data.

Assessment scale performed:

- National shared
- Regional
- Global

Approaches used:

Production model

Assessment purpose:

Regional stock assessments for Eastern Central Atlantic.

FAO Regional reviews for Eastern Central Atlantic.

reports/publications/data regularly PRODUCED:

name:	Sampling methods applied to fisheries science: a manual.
type:	technical guide
publishing frequency:	2005
url if any:	http://www.fao.org/docrep/009/a0198e/a0198e00.htm

name:	Report of the Sardine (Sardina pilchardus) Otolith Workshop.
type:	global assessment
publishing frequency:	2002
url if any:	http://www.fao.org/DOCREP/005/Y4097B/Y4097B00.HTM

name:	Report of the Workshop to plan the 1999 R/V DR. FRIDTJOF NANSEN surveys in the northern CECAF area and the standardization of acoustic surveys in the region.
type:	methodology
publishing frequency:	2001
url if any:	ftp://ftp.fao.org/docrep/fao/006/y2097e/y2097e00.pdf

name:	Report of the Workshop on the Small Pelagic Resources of Angola, Congo and Gabon.
type:	shared stock assessment
publishing frequency:	2000
url if any:	ftp://ftp.fao.org/docrep/fao/007/x8118e/x8118e00.pdf

name:	Report of the FAO/CECAF Working Group on the Assessment of Demersal Resources
type:	regional assessment

publishing frequency:	biannual
url if any:	ftp://ftp.fao.org/docrep/fao/009/a0536b/a0536b.zip

reports/publications/data regularly CONSUMED when producing:

name:	FIDI statistics
type:	time series
url if any:	http://www.fao.org/figis/servlet/static?dom=root&xml=tseries/index.xml
contact name:	Tsuji, Sachiko (FIDI)
contact email:	Sachiko.Tsuji@fao.org
contact phone:	+39 0657055318
contact organisation:	FAO/FIDI

name:	CECAF Small Pelagic Resources
type:	shared stock assessment
url if any:	ftp://ftp.fao.org/docrep/fao/007/x8118e/x8118e00.pdf
contact name:	Gamel Abdul Nasser Road, ACCRA, Ghana
contact email:	FAO-RAF@field.fao.org
contact phone:	+233 21 675 000/675051-060/701 0930
contact organisation:	FAO REGIONAL

name:	CECAF Assessment of Demersal Resources
type:	regional assessment
contact name:	Gamel Abdul Nasser Road, ACCRA, Ghana
contact email:	FAO-RAF@field.fao.org

contact phone:	+233 21 675 000/675051-060/701 0930
contact organisation:	FAO REGIONAL

Notes:

Working group reports for CECAF:

- Small pelagics, annual, NW Africa
- Demersal, bi-annual, by region
- Artisinal – more of a socio-economic report

Process:

scientists from nations who have national waters in CECAF region + scientists from countries who just fish there, e.g. spain. france, holland, gather together, each bringing their own data on catch and effort. These data may be national and/or may be from research insitutions and/or independend scientific cruise survey data. **This input data is private to the working group**.

There are subgroups for various stocks. Data for shared stocks is harmonised and grouped.

Generally a production model is used as data is often poor and does not allow for H-based or L-based models that require biologic data.

Data used for the production model is a mixture of CPUE from members, FAO data (formed from data submitted by member countries) and survey series. There are sometimes discrepancies between FAO data and country data though there shouldn't be as FAO reports what the countries give them.

The working group reports consist of catch, effort, trends, assessment. Also some quality assurance data. They are public.

These reports go to the scientific sub-committee, which then goes to committee and if adopted implies "voluntary compliance".

Though they would like to use more sophisticated models, data is poor. Some allowance is made if for example it is known that it was a “bad year” environmentally. They would like to incorporate data on upwelling.

They do not do ecosystem modelling, it is not part of their mandate but some projects have done such work.

For the regional reviews for FAO, Marete uses a combination of FIDI statistics and assessment reports.

Other Resources:

1. CSRP of Senegal has information on regional projects http://www.csrpsp.org/index.html
2. Morocco has good project information for NW africa: http://www.onp.co.ma/site_libre/projet/index.asp
3. Mauritania assessment every four years:

Evaluation des stocks et aménagement des pêcheries de la ZEE mauritanienne. Rapport du cinquième Groupe de travail IMROP. Nouadhibou, Mauritanie, 9-17 décembre 2002.
Failler, P.; Diop, M.; Dia, M.A. ; O/Inejih, C.A.; Tous, P. (éds).
CECAF/ECAF Series 06/662006

4. Large Marine Ecosystems : http://www.edc.uri.edu/lme/text/canary-current.htm
5. Oceanographic portal : http://www.iode.org/oceanportal/
6. Nansen research vessel for West Africa : http://www.cdcf.no/programmes/nansen/more
7. Guinea Current Large Marine Ecosystem Project: http://www.gclme.org/start.htm
8. Benguela: Predicting a Large Marine Ecosystem: http://www.bclme.org

Inventory Interview Form

name: Marcelo Vasconcellos

email: Marcelo.Vasconcellos@fao.org

webpage: CITES, WCPTC

areas of expertise: assessments, capacity building, technical assessment assistance

region: brazil, lesser antilles, nicaragua, belize, guatemala, papua new guinea, indonesia

Regional Fisheries Bodies: none

Standards: CITES, GIS

Notes:

Assessment:

Marcelo works often at national level, assisting national fisheries ministries in producing assessment data and/or reports. These tend to be countries in which either data is scant/low quality and/or technical capacity is low. In these cases he helps examine the available data to see what can be used, how it can be extrapolated, etc.

Data type is landing statistics and/or trade information (export statistics, volume and maybe value by species). There may be biologic data on size, age, maturity. There is almost never effort data as it is generally small-scale fisheries. In many cases FAO data must be used, and it is sometimes the case that data reported to FAO by the country does not match their own internal statistics. National statistics are not normally available on the web, and are often not public at all. Trade info is also often that which is reported to FAO. It may also be found on the CITES website.

A typical work flow would be the following:

1. Work with Papua in a technical and capacity building role to build a database of their raw data.
2. Raw data is passed to SPC regional fisheries body who develops assessments.
3. These may in turn be passed on to another body that makes management recommendations, such as WCPTC

CITES:

This is a convention that requires that if a species is on the CITES list of endangered species, then it cannot be exported without evidence of an assessment having been done that indicates that it is harvested at sustainable levels. Marcelo thus often works both with CITES to determine if a country's assessment is acceptable, and with countries to help them develop assessments for species on the CITES list.

Ecosystem approach

Marcelo is involved in several projects looking at ecosystem approaches. Papua is one example where they are interested in tuna and shark interaction. There is capacity building going on by for example improving monitoring of catches (example would be tuna longlining in which part of the catch is eaten by other fish before being hauled in, making it difficult to get the actual catch before depradation)

Also involved in an ecosystem approach in the caribbean looking at reef areas and getting an idea of the species distribution and interaction.

Reports created at FAO:

Some regional reviews

Technical papers on montoring and management

Case studies for technical projects

Types of data regularly used (see next page for details)

country data,

FIDI/FAO data,

Fishbase bio data for ecosystem modelling (growth, mortality data)

LANDSAT data (to estimate reef coverage, again for ecosystem approaches)

Various GIS data for depth, marine productivity

"Sea around us" for country level snapshots

reports/publications/data regularly consulted:

name	type	first pub	pub freq	url (if any)
FIDI statistics	time series	1950's	Annual	http://www.fao.org/figis/servlet/static?dom=root&xml=species/species_search.xml
Description: catch statistics for major commercial species				

name	type	first pub	pub freq	url (if any)
CITES	commodity	1975	Annual	http://www.unep-wcmc.org/citestrade/trade.cfm
Description: trade data for endangered species				

name	type	first pub	pub freq	url (if any)
Fishbase	bio data	?	?	http://www.fishbase.org/search.php
Description: FishBase on the web contains practically all fish species known to science. Mostly biological information.				

name	type	first pub	pub freq	url (if any)
Millennium coral reef	GIS	1999	?	http://imars.usf.edu/corals/
Description: mapping of global coral reefs using Landsat images from 1999-2002				

name	type	first pub	pub freq	url (if any)
UBC Fisheries Centre				http://www.seaaroundus.org/
Description: The data presented, which are all freely available, are meant to support studies of global fisheries trends and the development of sustainable, ecosystem-based fisheries policies.				

Inventory Interview Form

Name: John Beddington

Email: j.beddington@mrag.co.uk

Webpage: www.mrag.co.uk, www.imperial.ac.uk

NeOn-project.org

Areas of expertise: Population Dynamics/Stock Assessment

Regions/nations: Antarctic, South Atlantic, Indian Ocean, British Indian Ocean Territory (BIOT)

Regional fisheries bodies: Commission for the Conservation of Antarctic Living Marine Resources (CCAMLR), IOTC

Data facets worked with:

- Biologic data (species, age, mortality, length) Yes
- Catch data (landings, biomass, and exports) Yes
- Effort data (vessel type, vessel size, fishing gear type, crew size, trip duration, trip distance) Yes

Assessment scale performed:

- National For Falkland Islands, Inshore British Indian Ocean Territory
- Shared For South Georgia with CCAMLR
- Regional IOTC longline tuna
- Global

Approaches used:

Bayesian Modified De Lury, Mark Recapture, Management Strategy Evaluation, Age Structured Production Models. Age Structured Models tuned to CPUE and /or Mark Recapture Data.

Assessment purpose:

1. Falkland Islands, BIOT for regulation of Allowable Effort
2. CCAMLR for TACs
3. IOTC to inform Commission

reports/publications/data regularly PRODUCED:

name:	Papers for CCAMLR on Toothfish in South Georgia
type:	Scientific Paper
publishing frequency:	Annual
url if any:	CCAMLR Sciences
name:	Papers for IOTC
type:	Paper to working group meetings
publishing frequency:	Annual
url if any:	IOTC website

name:	Papers for Falkland Government
type:	Confidential Reports
publishing frequency:	2-3 times per year
url if any:	

name:	Papers for BIOT Government
type:	Confidential Reports
publishing frequency:	2-3 times per year
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

reports/publications/data regularly CONSUMED when producing:

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	
name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

Inventory Interview Form

Name: Dr Martin Genner

Email: m.genner@hull.ac.uk

Webpage: none

Areas of expertise: With respect to fisheries assessment, I have worked on long-term changes in demersal fish communities, relating these to environmental variables including regional climatic variability and regional fishing effort.

Regions/nations: Western English Channel, UK

Regional fisheries bodies: The Marine Fisheries Agency are responsible for enforcement (an agency of DEFRA).

CEFAS undertake survey work in the area.

Data facets worked with:

- Biologic data (species, age, mortality, length)

During MBA trawls, data collected are on fish species, individual length and individual weight

- Catch data (landings, biomass, and exports)

A regular survey trawl aimed at assessing fish species abundance (whole community), length distributions and biomass

- Effort data (vessel type, vessel size, fishing gear type, crew size, trip duration, trip distance)

Vessel type, size, fishing gear and trawl duration are recorded.

Assessment scale performed:

- National
- Shared
- Regional
- Global

Regional (Western English Channel)

Approaches used:

The demersal fish assemblage off Plymouth has been sampled intermittently between 1913 and 1986. This time series was initiated again in 2001, and has been running until present (Jan 2007). In total over 90 species have been recorded within 800+ otter trawls of approximately 1 hour each. The number of individual fish are counted, and length and weight recorded. Trawls are undertaken at 30-50 m depth over an area covering 42 x 19 km off Plymouth, but focusing on site L4 (50°15.5′ N, 04°13′ W). During the time-series seven vessels were used, ranging in overall length from 18.3 to 39.0m. The trawl gears have been of comparable dimensions, and trawling has been carried out at similar speeds.

Assessment purpose:

These long-term dataset has been used to identify long-term trends in the responses of fish to fluctuations in climate, and to human-induced changes associated with commercial exploitation of stocks.

reports/publications/data regularly PRODUCED:

The research has been published in peer-reviewed journals, rather than regular reports.

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

name:	
type:	
publishing frequency:	
url if any:	

reports/publications/data regularly CONSUMED when producing:

Primarily the current primary research literature, rather than regular reports in 'grey' literature.

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

name:	
type:	
url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

name:	
type:	

url if any:	
contact name:	
contact email:	
contact phone:	
contact organisation:	

Notes

Annex II – Relative importance of system descriptors

TITLE	RATING 1=critical
4.1. General aspects	
4.1.1. Selection criteria	
Why was the system selected for inclusion?	3
Which kind of experts uses the system in their work?	3
Who are the typical users of the system and how do they use the system?	3
In what broad ways does it satisfy the needs of the case study?	3
In what broad ways does it satisfy the needs of NeOn project partners?	3
Is the system authoritative?	3
Is the system encyclopedic?	3
4.1.2. System relations criteria	
How does this system relate to other systems?	2
Are there applications that use this system together with other systems to generate a synthetic output?	2
Is there overlap between this system and other systems in the inventory?	3
4.2. Data aspects	
4.2.1. Data	
Coverage	1
Resource Type	1
4.2.2. Metadata	
What kind of classification system?	2
Maintained by whom?	2
How updated?	2
How widely used?	3
How expressed?	3
4.2.3. Quality assurance data	
Where does data come from?	2
How is data collected?	2
How frequently?	2
For how long?	2
By whom?	2
According to what protocols?	2
How verified? Random re-sampling?	2
How processed and handled?	2
How is it input into an information system?	2
How is it verified? Double entry?	2
Who stands behind the data validity?	1
What is their authority?	1
4.2.4. Data structure	1
4.3. Availabilty aspects	
4.3.1. Access	
Is the data easily accessible/available?	1

TITLE	RATING 1=critical
What are the access rights?	1
4.3.2. Copyright	
What are the rules if any for attributing ownership?	1
What are the rules if any for citation?	1
4.4. Technical aspects	
4.4.1. How can the data be accessed or exposed?	1
4.4.2. What are the APIs?	
Are they published?	1
Are they stable?	1
How is change handled?	2
Are they extensible?	1
Are they documented?	1
4.4.3. How efficient is data retrieval?	
What kind of load can the system handle?	3
How often does the structure of the data change?	1
4.5. Administrative aspects	
4.5.1. Who has authority for the system?	1
4.5.2. Who can change the data?	2
4.5.3. Who can modify the system?	2
4.6. Practical aspects	
Are they interested in working with the NeOn case study?	1
Can they grant access to the data?	1
Any restrictions or conditions?	1
Can they modify or extend the API if needed?	1
Can they modify the data structure if needed?	1
Do they have resources to expend for interaction with NeOn?	1
Do they have resources to expend were there a need for API modification/extension?	1

Figure 4. Relative importance of system descriptors

Annex III – XML schema representation of system descriptors

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
attributeFormDefault="unqualified">
	<xs:element name="accessRights">
		<xs:annotation>
			<xs:documentation>what are the rules regarding access and use of the
data</xs:documentation>
		</xs:annotation>
	</xs:element>
	<xs:element name="acronym">
		<xs:annotation>
			<xs:documentation>an abbreviated name for a formal title</xs:documentation>
		</xs:annotation>
	</xs:element>
	<xs:element name="administrative">
		<xs:annotation>
			<xs:documentation> administrative data about a system</xs:documentation>
		</xs:annotation>
		<xs:complexType>
			<xs:sequence>
				<xs:element ref="authority"/>
				<xs:element ref="changeAuthority"/>
			</xs:sequence>
		</xs:complexType>
	</xs:element>
	<xs:element name="applicationProfile">
		<xs:annotation>
			<xs:documentation>a formal system of concepts used to describe
data</xs:documentation>
		</xs:annotation>
		<xs:complexType>
			<xs:sequence>
				<xs:element ref="name"/>
				<xs:element ref="url"/>
			</xs:sequence>
		</xs:complexType>
	</xs:element>
	<xs:element name="authority">
		<xs:annotation>
			<xs:documentation>who has authority for the system</xs:documentation>
		</xs:annotation>
		<xs:complexType>
			<xs:sequence>
				<xs:element ref="contact"/>
			</xs:sequence>
		</xs:complexType>
	</xs:element>
	<xs:element name="changeAuthority">
		<xs:annotation>
			<xs:documentation>who has authority to change the system</xs:documentation>
		</xs:annotation>
```

```
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="contact"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="citation">
                <xs:annotation>
                        <xs:documentation>rules for citing data</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="classificationSystem">
                <xs:annotation>
                        <xs:documentation>an formal system of concepts used to organise
data</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="collection">
                <xs:annotation>
                        <xs:documentation>how is the data collected and maintained or
updated</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="method"/>
                                <xs:element ref="frequency"/>
                                <xs:element ref="source"/>
                                <xs:element ref="collectionProtocol"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="collectionProtocol">
                <xs:annotation>
                        <xs:documentation>a formally defined method used to collect
data</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="conditions">
                <xs:annotation>
                        <xs:documentation>any conditions or restrictions</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true">
                        <xs:attribute name="canModifyAPI" type="xs:boolean" use="required"/>
                        <xs:attribute name="canModifyStructure" type="xs:boolean" use="required"/>
                </xs:complexType>
        </xs:element>
        <xs:element name="contact">
                <xs:annotation>
                        <xs:documentation> information about how to contact system
owners</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="name"/>
                                <xs:element ref="email"/>
                                <xs:element ref="telephone"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="content">
```

```
                <xs:annotation>
                        <xs:documentation>what the service provides</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:choice maxOccurs="unbounded">
                                <xs:element ref="data"/>
                        </xs:choice>
                </xs:complexType>
        </xs:element>
        <xs:element name="copyright">
                <xs:annotation>
                        <xs:documentation>how should the data be credited</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="data">
                <xs:annotation>
                        <xs:documentation>the data of interest</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="name"/>
                                <xs:element ref="description"/>
                                <xs:element ref="owner"/>
                                <xs:element ref="spatialCoverage"/>
                                <xs:element ref="temporalCoverage"/>
                                <xs:element ref="languageCoverage"/>
                                <xs:element ref="type"/>
                                <xs:element ref="mime-type" maxOccurs="unbounded"/>
                                <xs:element ref="encoding"/>
                                <xs:element ref="structure"/>
                                <xs:element ref="keys"/>
                                <xs:element ref="metadata"/>
                                <xs:element ref="collection"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="description">
                <xs:annotation>
                        <xs:documentation> an overview of an item</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="documentation">
                <xs:annotation>
                        <xs:documentation>the URL of any supporting documentation</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="email">
                <xs:annotation>
                        <xs:documentation>email address</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="encoding">
                <xs:annotation>
                        <xs:documentation>the data encoding used, such as UTF-8 or
Binary</xs:documentation>
                </xs:annotation>
                <xs:simpleType>
                        <xs:restriction base="xs:token">
                                <xs:enumeration value="UTF-8"/>
                                <xs:enumeration value="ISO 8859-1"/>
```

```
                        <xs:enumeration value="ISO 8859-2"/>
                        <xs:enumeration value="ASCII"/>
                        <xs:enumeration value=""/>
                    </xs:restriction>
                </xs:simpleType>
            </xs:element>
            <xs:element name="encodingScheme">
                <xs:annotation>
                    <xs:documentation>a formal method of data notation</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="example">
                <xs:annotation>
                    <xs:documentation>an demonstrative instance of something
described</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="frequency">
                <xs:annotation>
                    <xs:documentation>when is data updated</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="gathering">
                <xs:annotation>
                    <xs:documentation>how is data gathered</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="general">
                <xs:annotation>
                    <xs:documentation>descriptive information about a system</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                    <xs:sequence>
                        <xs:element ref="name"/>
                        <xs:element ref="acronym"/>
                        <xs:element ref="systemType"/>
                        <xs:element ref="description"/>
                        <xs:element ref="url"/>
                        <xs:element ref="contact"/>
                        <xs:element ref="overlap"/>
                        <xs:element ref="selectionCriteria"/>
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
            <xs:element name="handling">
                <xs:annotation>
                    <xs:documentation>how is data handled</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="humanResources">
                <xs:annotation>
                    <xs:documentation>do they have time to work with NeOn</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true">
                    <xs:attribute name="canExpendResources" type="xs:boolean" use="required"/>
                </xs:complexType>
            </xs:element>
            <xs:element name="id">
                <xs:annotation>
                    <xs:documentation>a unique ID value</xs:documentation>
```

```
            </xs:annotation>
        </xs:element>
        <xs:element name="input">
            <xs:annotation>
                <xs:documentation>how is data input</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="interest">
            <xs:annotation>
                <xs:documentation>are they interested in working with NeOn</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="interface">
            <xs:annotation>
                <xs:documentation>describes a method by which a service may be
accessed</xs:documentation>
            </xs:annotation>
            <xs:complexType>
                <xs:sequence>
                    <xs:element ref="name"/>
                    <xs:element ref="description"/>
                    <xs:element ref="example"/>
                    <xs:element ref="location"/>
                    <xs:element ref="protocol"/>
                    <xs:element ref="mime-type"/>
                    <xs:element ref="encoding"/>
                    <xs:element ref="structure"/>
                    <xs:element ref="parameters"/>
                    <xs:element ref="documentation"/>
                    <xs:element ref="loadHandling"/>
                </xs:sequence>
                <xs:attribute name="published" type="xs:boolean" use="required"/>
                <xs:attribute name="stable" type="xs:boolean" use="required"/>
                <xs:attribute name="extensible" type="xs:boolean" use="required"/>
                <xs:attribute name="boolean" type="xs:boolean" use="required"/>
            </xs:complexType>
        </xs:element>
        <xs:element name="key">
            <xs:annotation>
                <xs:documentation>time range the data covers</xs:documentation>
            </xs:annotation>
            <xs:complexType mixed="true">
                <xs:sequence>
                    <xs:element ref="name"/>
                    <xs:element ref="classificationSystem" maxOccurs="unbounded"/>
                </xs:sequence>
            </xs:complexType>
        </xs:element>
        <xs:element name="keys">
            <xs:annotation>
                <xs:documentation>time range the data covers</xs:documentation>
            </xs:annotation>
            <xs:complexType mixed="true">
                <xs:choice maxOccurs="unbounded">
                    <xs:element ref="key"/>
                </xs:choice>
            </xs:complexType>
        </xs:element>
        <xs:element name="language">
            <xs:annotation>
```

```
                        <xs:documentation>an ISO-2 language code</xs:documentation>
                    </xs:annotation>
                </xs:element>
                <xs:element name="languageCoverage">
                    <xs:annotation>
                        <xs:documentation>the languages in which some or all of the resource can be
had</xs:documentation>
                    </xs:annotation>
                    <xs:complexType>
                        <xs:sequence>
                            <xs:element ref="language" maxOccurs="unbounded"/>
                        </xs:sequence>
                    </xs:complexType>
                </xs:element>
                <xs:element name="legal">
                    <xs:annotation>
                        <xs:documentation> legal data about the use of a system or its
data</xs:documentation>
                    </xs:annotation>
                    <xs:complexType>
                        <xs:sequence>
                            <xs:element ref="accessRights"/>
                            <xs:element ref="copyright"/>
                            <xs:element ref="citation"/>
                        </xs:sequence>
                    </xs:complexType>
                </xs:element>
                <xs:element name="loadHandling">
                    <xs:annotation>
                        <xs:documentation>how much access can the interface
handle?</xs:documentation>
                    </xs:annotation>
                </xs:element>
                <xs:element name="location">
                    <xs:annotation>
                        <xs:documentation>where an interface can be accessed</xs:documentation>
                    </xs:annotation>
                </xs:element>
                <xs:element name="method">
                    <xs:annotation>
                        <xs:documentation>how is data collected</xs:documentation>
                    </xs:annotation>
                    <xs:complexType>
                        <xs:sequence>
                            <xs:element ref="gathering"/>
                            <xs:element ref="handling"/>
                            <xs:element ref="input"/>
                            <xs:element ref="validation"/>
                        </xs:sequence>
                    </xs:complexType>
                </xs:element>
                <xs:element name="metadata">
                    <xs:annotation>
                        <xs:documentation>describes the data</xs:documentation>
                    </xs:annotation>
                    <xs:complexType mixed="true">
                        <xs:sequence maxOccurs="unbounded">
                            <xs:element ref="encodingScheme" maxOccurs="unbounded"/>
                            <xs:element ref="applicationProfile" maxOccurs="unbounded"/>
                        </xs:sequence>
```

```
            </xs:complexType>
        </xs:element>
        <xs:element name="mime-type">
            <xs:annotation>
                <xs:documentation>describes the data mime type</xs:documentation>
            </xs:annotation>
            <xs:simpleType>
                <xs:restriction base="xs:token">
                    <xs:enumeration value="text/xml"/>
                    <xs:enumeration value="text/html"/>
                    <xs:enumeration value="text/plain"/>
                    <xs:enumeration value="text/richtext"/>
                    <xs:enumeration value="text/tab-separated-values"/>
                    <xs:enumeration value="text/csv"/>
                    <xs:enumeration value="text/sgml"/>
                    <xs:enumeration value="application/postscript"/>
                    <xs:enumeration value="application/rtf"/>
                    <xs:enumeration value="application/pdf"/>
                    <xs:enumeration value="application/msaccess"/>
                    <xs:enumeration value="application/msword"/>
                    <xs:enumeration value="application/sgml"/>
                    <xs:enumeration value="application/x-framemaker"/>
                    <xs:enumeration value="application/octet-stream"/>
                    <xs:enumeration value="application/mspowerpoint"/>
                    <xs:enumeration value="application/ms-excel"/>
                    <xs:enumeration value=""/>
                </xs:restriction>
            </xs:simpleType>
        </xs:element>
        <xs:element name="name">
            <xs:annotation>
                <xs:documentation>a formal title</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="NeOn">
            <xs:annotation>
                <xs:documentation> information of relevance to NeOn for a
system</xs:documentation>
            </xs:annotation>
            <xs:complexType>
                <xs:sequence>
                    <xs:element ref="interest"/>
                    <xs:element ref="conditions"/>
                    <xs:element ref="humanResources"/>
                </xs:sequence>
            </xs:complexType>
        </xs:element>
        <xs:element name="overlap">
            <xs:annotation>
                <xs:documentation>overlap with other systems in inventory</xs:documentation>
            </xs:annotation>
            <xs:complexType mixed="true">
                <xs:choice minOccurs="0" maxOccurs="unbounded">
                    <xs:element ref="systemLink"/>
                </xs:choice>
            </xs:complexType>
        </xs:element>
        <xs:element name="owner">
            <xs:annotation>
                <xs:documentation>the owner of a resource</xs:documentation>
```

```
                </xs:annotation>
        </xs:element>
        <xs:element name="parameter">
                <xs:annotation>
                        <xs:documentation> a parameter needed to access a result from a
service</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="name"/>
                                <xs:element ref="description"/>
                                <xs:element ref="values"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="parameters">
                <xs:annotation>
                        <xs:documentation>parameters needed to access a result from a
service</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:choice maxOccurs="unbounded">
                                <xs:element ref="parameter"/>
                        </xs:choice>
                </xs:complexType>
        </xs:element>
        <xs:element name="protocol">
                <xs:annotation>
                        <xs:documentation>the specific terms for communicating with an
interface</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="relations">
                <xs:annotation>
                        <xs:documentation>relations to other systems</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="subSystems"/>
                                <xs:element ref="superSystems"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="relationToCaseStudy">
                <xs:annotation>
                        <xs:documentation>why was the system selected for inclusion</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="relationToNeOn">
                <xs:annotation>
                        <xs:documentation>why was the system selected for inclusion</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="selectionCriteria">
                <xs:annotation>
                        <xs:documentation>criteria of system selected inclusion</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="userTypes"/>
```

```
                        <xs:element ref="relationToCaseStudy"/>
                        <xs:element ref="relationToNeOn"/>
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
            <xs:element name="service">
                <xs:annotation>
                    <xs:documentation>a component of a system that provides a specific service such
as data exchange</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                    <xs:choice>
                        <xs:element ref="interface" maxOccurs="unbounded"/>
                    </xs:choice>
                </xs:complexType>
            </xs:element>
            <xs:element name="source">
                <xs:annotation>
                    <xs:documentation>where does the data come from?</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element name="spatialCoverage">
                <xs:annotation>
                    <xs:documentation>geo area the data covers</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true"/>
            </xs:element>
            <xs:element name="structure">
                <xs:annotation>
                    <xs:documentation>describes the organisational paradigm for the way data is
stored</xs:documentation>
                </xs:annotation>
                <xs:simpleType>
                    <xs:restriction base="xs:token">
                        <xs:enumeration value="relational"/>
                        <xs:enumeration value="hierarchical"/>
                        <xs:enumeration value="unstructured"/>
                        <xs:enumeration value="ontological"/>
                        <xs:enumeration value="object-oriented"/>
                        <xs:enumeration value="flat"/>
                        <xs:enumeration value=""/>
                    </xs:restriction>
                </xs:simpleType>
            </xs:element>
            <xs:element name="subSystems">
                <xs:annotation>
                    <xs:documentation>systems used by the system to be described in another
document</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true">
                    <xs:choice minOccurs="0" maxOccurs="unbounded">
                        <xs:element ref="systemLink"/>
                    </xs:choice>
                </xs:complexType>
            </xs:element>
            <xs:element name="superSystems">
                <xs:annotation>
                    <xs:documentation>systems that use this system</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true">
```

```
                        <xs:choice minOccurs="0" maxOccurs="unbounded">
                                <xs:element ref="systemLink"/>
                        </xs:choice>
                </xs:complexType>
        </xs:element>
        <xs:element name="system">
                <xs:annotation>
                        <xs:documentation>The information system being described. This is the top-level
element.</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="general"/>
                                <xs:element ref="NeOn"/>
                                <xs:element ref="legal"/>
                                <xs:element ref="administrative"/>
                                <xs:element ref="content"/>
                                <xs:element ref="service"/>
                                <xs:element ref="relations"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="systemLink">
                <xs:annotation>
                        <xs:documentation>a reference to a system described
elsewhere</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:attribute name="idref" use="required"/>
                </xs:complexType>
        </xs:element>
        <xs:element name="systemType">
                <xs:annotation>
                        <xs:documentation>what kind of system</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:attribute name="authoritative" type="xs:boolean" use="required"/>
                        <xs:attribute name="encyclopaedic" type="xs:boolean" use="required"/>
                </xs:complexType>
        </xs:element>
        <xs:element name="temporalCoverage">
                <xs:annotation>
                        <xs:documentation>time range the data covers</xs:documentation>
                </xs:annotation>
                <xs:complexType mixed="true"/>
        </xs:element>
        <xs:element name="telephone">
                <xs:annotation>
                        <xs:documentation>telephone number</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="type">
                <xs:annotation>
                        <xs:documentation>describes the data type, e.g. webpage, document, time-
series</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="url">
                <xs:annotation>
```

```
                        <xs:documentation> how to access some web presence of a system, or a query for
a service</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="userTypes">
                <xs:annotation>
                        <xs:documentation>what kind of users the system has, e.g. researchers, public,
managers</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="validation">
                <xs:annotation>
                        <xs:documentation>how is data validated</xs:documentation>
                </xs:annotation>
        </xs:element>
        <xs:element name="value">
                <xs:annotation>
                        <xs:documentation>a value</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                        <xs:sequence>
                                <xs:element ref="id"/>
                                <xs:element ref="description"/>
                        </xs:sequence>
                </xs:complexType>
        </xs:element>
        <xs:element name="values">
                <xs:annotation>
                        <xs:documentation>a range of values</xs:documentation>
                </xs:annotation>
        </xs:element>
</xs:schema>
```

Annex IV – List of Acronyms used in this deliverable

AGROVOC	FAO-Multilingual Agricultural Thesaurus
API	Application Programming Interface
ASFA	Aquatic Sciences Fisheries Abstracts
CITES	Convention on International Trade in Endangered Species
CORBA	Common Object Remote Broker Architecture
CPUE	Fishery Basic Catch Per Unit Effort
CSV	Comma Separated Value files
EIMS	Electronic Information Management System
FAOLEX	FAO-Legislative database
FI	Fisheries Department, FAO of the UN
FIDI	Fishery Information, Data and Statistics Unit, FAO of the UN
FIGIS	Fisheries Global Information System
FIRMS	Fishery Resources Monitoring System
FSDAS	Fisheries Stock Depletion Assessment System
GIS	Geographical Information Systems
HTML	HyperText Markup Language
ISO	International Organization for Standardization
ITIS	Integrated Taxonomic Information System
KIMS/KIDS	Key Indicator Mapping System/Key Indicator Data System
MIME	Multipurpose Internet Mail Extensions
MIT	Massachusetts Institute of Technology
NMFS	National marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
OBIS	Ocean Bio-geographic Information System
OO	Object Oriented

OWL	Web Ontology Language
PDF	Portable Document Format
RDF	Resource Description Framework
RMI	Remote Method Invocation
RSS	Really Simple Syndication
RTMS	Fisheries Reference Table Management System
SGML	Standard Generalized Markup Language
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
UTF8	8 bit Unicode Transformation Format
WP	Work package
XML	Extensible Markup Language
XHTML	EXtensible HyperText Markup Language
XSL	EXtensible Stylesheet Language