How SeaDataNet has influenced data management methods in the IEO_DC improvements-benefits, changes, what more could be done

María Jesús García, Elena Tel

IEO Data Center

SDN2. Final plenary meeting Brest, 16,17 September 2015

Content

- IEO-DC Background

- SeaDataNet project
  - *improvements-benefits:*
    - ✓ *metadata*
    - ✓ *data types*
    - ✓ *data format & vocabularies*
    - ✓ *developing data management system*
    - ✓ *installations of tools for data communications*
    - ✓ *data processing; input data & quality of data and product*
  - *more to be done by:*
    - ✓ *the IEODC*
    - ✓ *the consortium*

IEO-DC Background

- *Period 1967-1978 (no computer at IEO)*
  - *the DC is stablish in the frame of the NODCs under IOC*
  - *the data were sending in preformated sheet or punched card to the WODC-A (USA) for proccesing the data an archiving.*
  - *the WODC sent regularly or under request data from the Spanish area by post mail in tapes*
- *Period 1979-1993 (starting to develop software)*
  - *developing software for cruise stations data proccesing and later on for time series*
- *Period 1994-2005 (working in European Data Projects)*
  - *Eurodimin, Medatlas, Medar, Sea-Search, Seadatanet.*

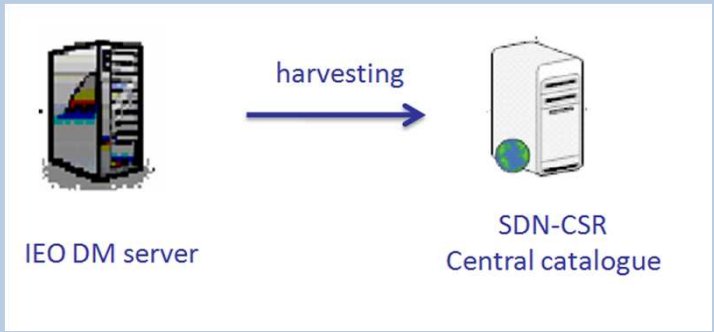
Improvements & benefits : create metadata

Before SDN	Improvements SDN	Benefits
•EDMO: No action	•EDMO: Created on line in the central catalogue (MARIS)	• Archiving and dissemination of the relevant information included in this central catalogue. • Advanced technical facilities for creation metadata avoiding errors. • Generating in automatic way means that the DC should have the information organized in files or RDB.
•EDMEDPR: No Action	•EDMEDPR: Created on line in the central catalogue (MARIS). We will change to generate through MIKADO .	
•EDMED: created in paper in the frame of Eurodimin project.	•EDMED: Created manually with MIKADO. We will change to generate automatically through MIKADO	
•EDIOS: created in paper in the frame of GOOS.	•EDIOS: created manually with MIKADO. We will change to generate automatically through MIKADO A	

Improvements & benefits: create metadata

Before SDN	Improvements SDN	Benefits
● Header Data Records: including header records at the first lines of the files. ● CSR(ROSCOP Form) ✓ Created by scientist after the cruise in paper or in digital form using an ICES tool (ROSWIN).  ICES CRUISE SUMMARY REPORT (ROSCOP) Centre: Originators Ref. No. Is data exchange restricted Yes In part No SHIP enter the full name and international radio call sign of the ship from which the data were collected, and indicate the type of ship, for example, research ship, ship of opportunity, naval survey vessel, etc. Name: Cornide de Saavedra Call Sign: ??? Type of ship: Research Vessel CRUISE NO. / NAME BOCADEVA0608 enter the unique number, name or acronym assigned to the cruise (or cruise leg, if appropriate). CRUISE PERIOD start (to/sail) 06/11/2010 day/month/year to 19/11/2010 day/month/year end (return to port) PORT OF DEPARTURE (enter name and country) Cádiz, Spain PORT OF RETURN (enter name and country) Cádiz, Spain RESPONSIBLE LABORATORY enter name and address of the laboratory responsible for coordinating the scientific planning of the cruise Name: Instituto Español de Oceanografía Address: Muelle de Levante s/n, Puerto Pesquero E-11006 Cádiz, Spain Country: Spain CHIEF SCIENTIST(S) enter name and laboratory of the person(s) in charge of the scientific work (chief of mission) during the cruise Ignacio Sobrino Yraola, Instituto Español de Oceanografía OBJECTIVES AND BRIEF NARRATIVE OF CRUISE enter sufficient information about the purpose and nature of the cruise so as to provide the context in which the report data were collected ✓ Sending to the DC for ✓ validate and prepare in digital form, if necessary.	•CDIs: Generation automatically the Common Data Index (CDIs) from the header records in the data files or from the RDB using MIKADO tools. ● CSRs ✓At the beginning of the project the method was the same but sending the CSRs to SDN central catalogue (BSH). ✓ Later on, at the DC the CSRs are created on line or using MIKADO tool. ✓Nowadays, the CSR are created on ship using MIKADO and given the reference according to the code ship and date of the cruise (unique). 	1) The CDIs are the bases for the data dissemination through the SDN data portal. 2) The use of MIKADO Advanced technical facilities for creation CSRs avoiding many errors and including other related metadata. 3) The Creation on ship improve the control of the unified references between CSR and the different data types register at the cruise and ensure the creation of the CSR.

Improvements & benefits: transfer metadata

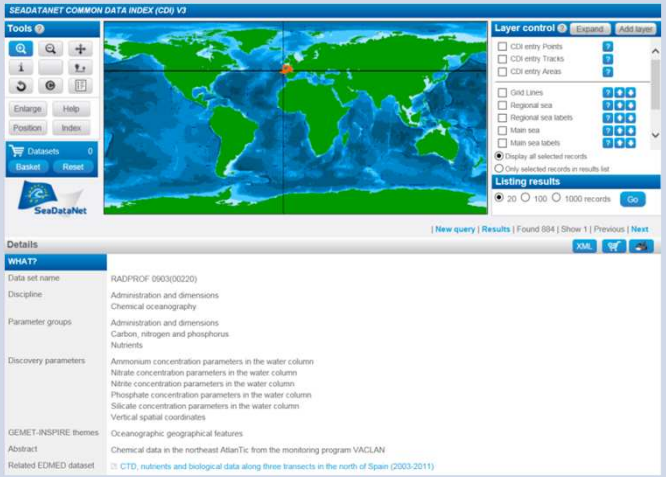
Before SDN	Improvements SDN	Benefits
• **EDMED: sending by post mail to the corresponding catalogue.** • **CSR and EDIOS: sending by e-mail to the corresponding catalogues.**	•All the metadata created by MIKADO was •Sending by e-mail to the corresponding catalogues   Same for EDMED and EDIOS catalogues ➢Nowadays for CSRs and CDIs by harvesting using GeoNetWork . 	• The e-mail is a common methods used for transferring information. • The harvesting is and advanced technical facilities to automate the METADATA transfer and improve the control and changes to be done on those metadata by the DC.

Improvements & benefits : METADATA

General Benefits for the DC & user in general

- Mikado tools have benn very helpul to create all the metadata records
- Relevant information from:

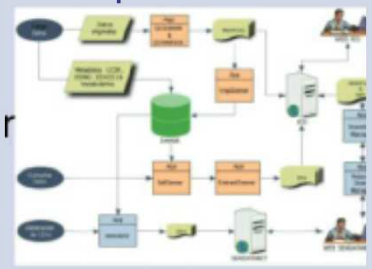
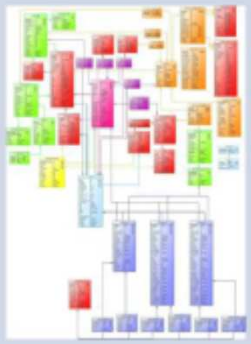
1) The new metadata catalogues: CDI, EDMEDPR, EDMO,
2) The new metadata Portals: EDMED,CDI,EDMEDPR, EDMO
3) The advance searching creteria in EDIOS & CSR
4) The related metadata Portals: CDI, CSR, EDMED,EDMEDPR,EDMO, EDIOS



CDI and the related EDMED

CSR and related Project

Developing data system & installation tools

Before SDN	Improvements SDN	Benefits
• RDB: only for cruise station in MySQL some selection criteria and level interpolation ➢  ➢Installation: on action	• RDB for many data types using SDN vocab. The design includes catalogues tables but that part is nor operational and should be designed according the SDN description   • Installation of different versions of SDN-DM • Installation of GeoNetWork	• Besides to include all data types and responsability of DC, the use of SDN vocab. and other standars (QCflags, etc) facilites the SDN work. • The DM allows to selec and download data from the DC and integrated. • The GeoNeWork allows the automatic data transfer from DC to central catalogues

Improvements & benefits : data types

Before SDN	Improvements SDN	Benefits
•Phisical, biochemical & meteorological data : ➢profiles registered by CTD, discrete water samplers (biochemical data nutrients, oxygen, primary production, etc). ➢time series registered by: tide gauges, current meters , thermistors chain , bouys and meteorological stations	•It has further been including : ➢Underway Data register by Thermosalinometer and fluorimeter. ➢Current profiler. ADCP ➢Pollutants And can be included bathymetry and biological data (egs& larvae), and we would like to include plankton acording to thecdescription done in SDN2,	compiling and processing more data types

Improvements & benefits: vocabularies & data format

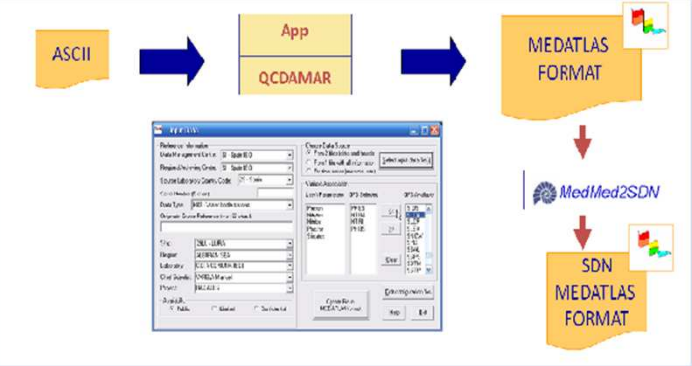
Before SDN	Improvements SDN	Benefits
•Vocabularies: Extended GF3 (web IFREMER), ROSCOP codes for data types & ICES codes for ships.	• Vocabularies: SDN vocabularies & libraries. (BODC web page) *For Pollutant data the IEO ask to BODC for including many parameters in the P01 vocabulary.*	• Standardization of parameters & library. • More possibilities to chose the data input depending on the data type.
•MEDATLAS Format	• SDN MEDATLAS Format • SDN ODV format for TG, TS & Pollutants • CFPOINT(NetCDF) for the moment just conversion from other format.	

Improvements & benefits : data proccesing(input)

Starting Methods

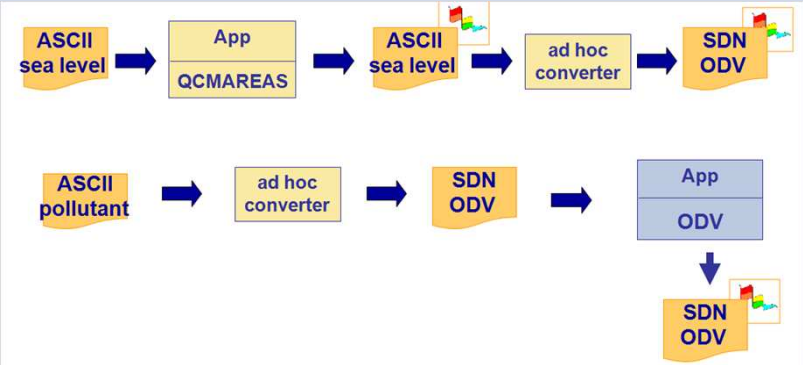
- SDN_Medatlas

1) QCDAMAR (input, & QC).
2) Med2MedSDN tool for including semantic lines,which include mapping vocabularies

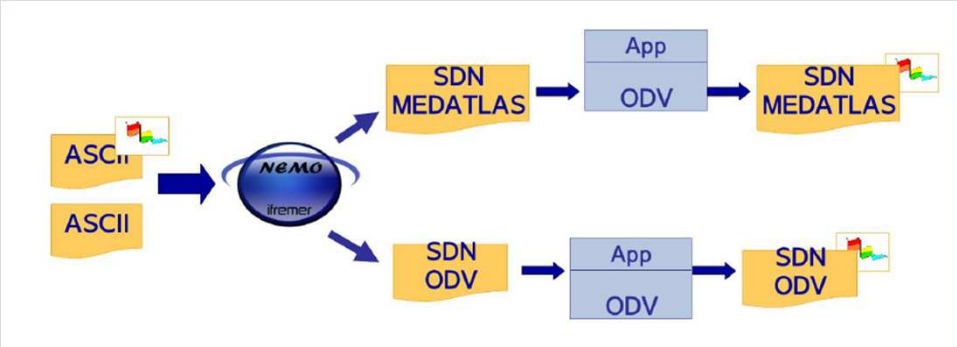


- ODV for TS, TG and pollutant

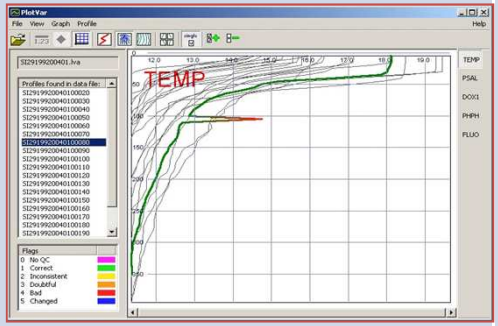
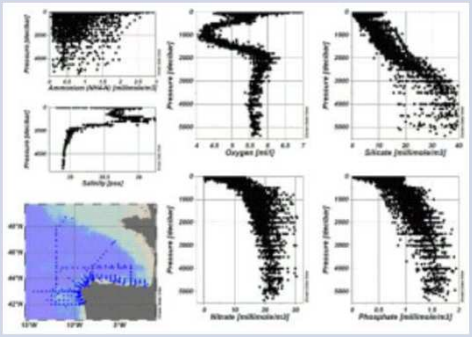
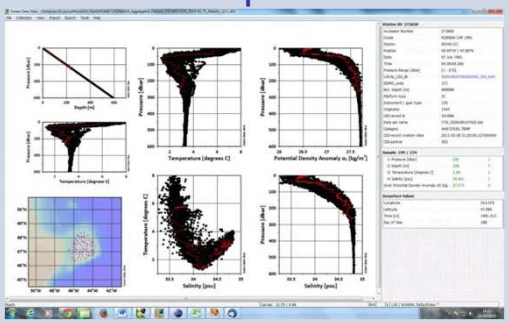
1) Preparing mapping and semantic lines
1) Ad_hoc converter tool



Improvements & benefits : data proccesing(input)

Improvements SDN Methods	Benefits
Nowadays: • SDN MEDATLAS: NEMO tool that include the semantic lines. 	Started to work with the SDN tools for two reason: 1) avoid errors in data format and ensure the SDN vocabulary 2) not needed to upgrade the QCDAMAR software
•The provider start to sending the data in MEDATLAS or ODV format using NEMO.	3) Beside facilitate the work to the DC team , the data can be better qualified due to a more easy and fluency communication having the same format and figures.

Improvements & benefits : Quality Data & Products

Before SDN	Improvements SDN	Benefits
QCDAMAR tools using MEDATLAS QC protocols (similar to SDN protocols). 	QCDAMAR cruise by cruise and ODV for a set of cruises of a project or a zone.  Example of dataset from radprof cruises used for QC between DC and provider.  Example of Fletan cruises dataset used during control assessment between the leader product and DC.	With NEMO and ODV we ensure better QC because we can have a better communication with the leader of the data product and with the provider through those tools.

General benefits & More to be done

- *Benefits :*

✓ *Standard tools for creating metadata, data and download data from all the DC*
✓ *Vocabularies for data standardization*
✓ *Data portal for data dissemination*
✓ *Related catalogues portal for having more complete and integrated information*
✓ *Improving data quality an product due continuous upgrading developing data analysis and higher volume of data.*

- *More to be done by:*
  - ✓ *the IEO: upgrading RDB for metadata and for more data type as organism.*
  - ✓ *the consortium: analyse the possibility to develop a metadata RDB for DC or upgrading Mikado to work similar to a metadata data base.*
  - ✓ *advance in automatate the catalogue transfer by using GeoNetWork or other tool.*

Thank you