WP9.1: Upgrading the Common Data Index (CDI) Data Discovery & Access service

introduction

By
Dick M.A. Schaap – MARIS
Technical Coordinator SeaDataCloud

Riga – Latvia, 30 November 2016,
SeaDataCloud CG meeting no 1

Common Data Index (CDI) service

- It provides users a highly detailed insight and unified access to the large volumes of marine and oceanographic data sets managed by the distributed data centres

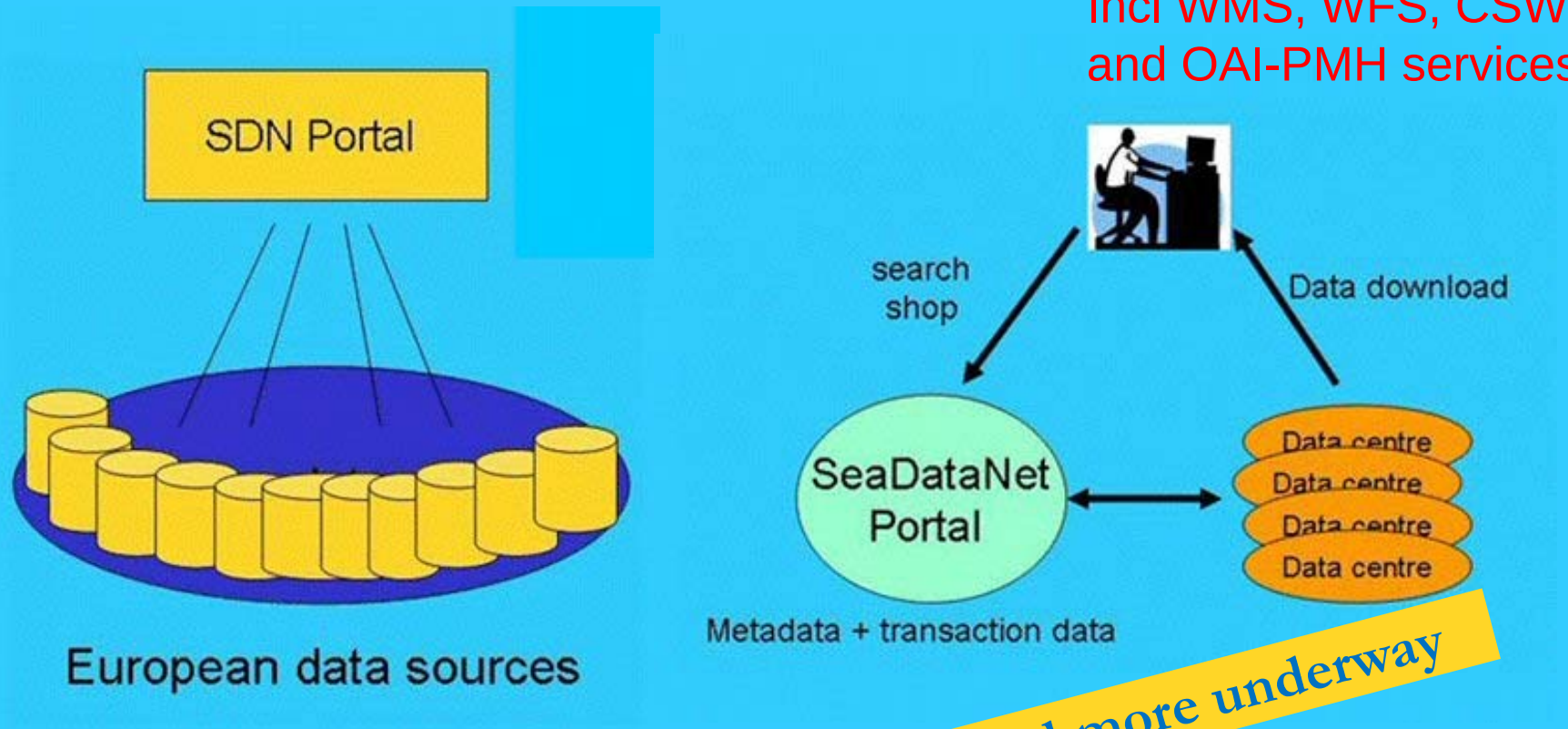
- It is a fine-grained index (ISO 19115 – ISO 19139) to individual data measurements (such as a CTD cast or moored instrument record).

- INSPIRE compliant and supported by Controlled Vocabularies, and Directories (EDMO, EDMERP, CSR, EDMED)

- MIKADO editor for preparation of metadata records and NEMO software for preparation of SeaDataNet ODV files

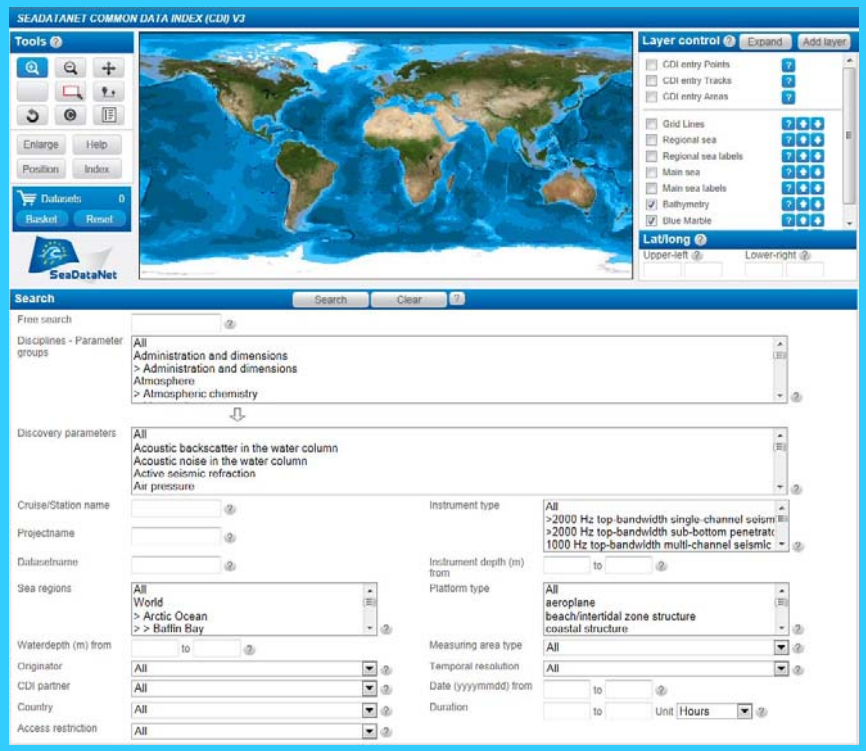
Common Data Index (CDI) service for discovery and unified access of data

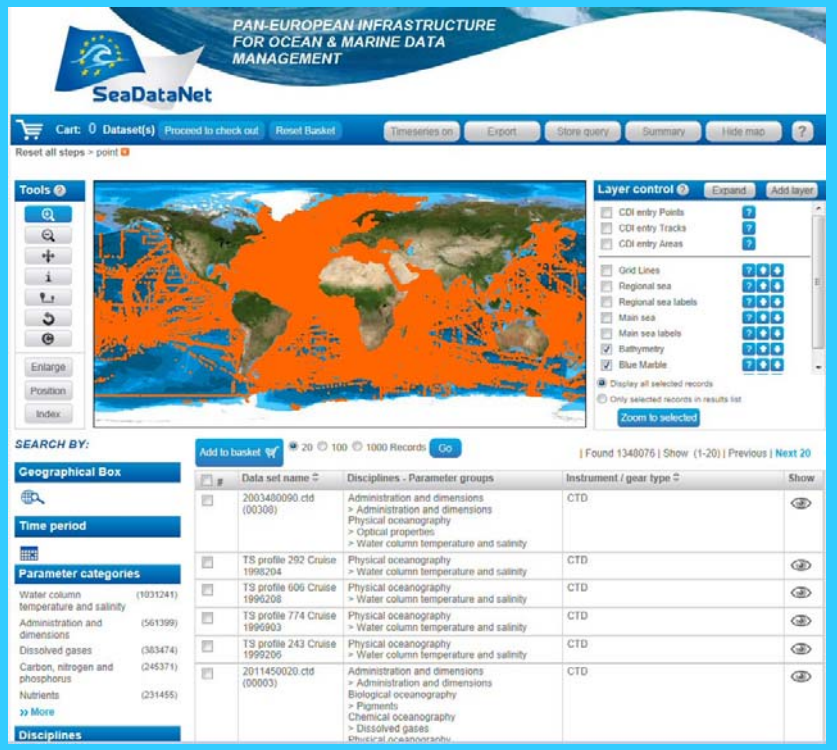
Incl WMS, WFS, CSW and OAI-PMH services



Already > 100 data centres connected and more underway

User interfaces for discovery and shopping





Extended Search

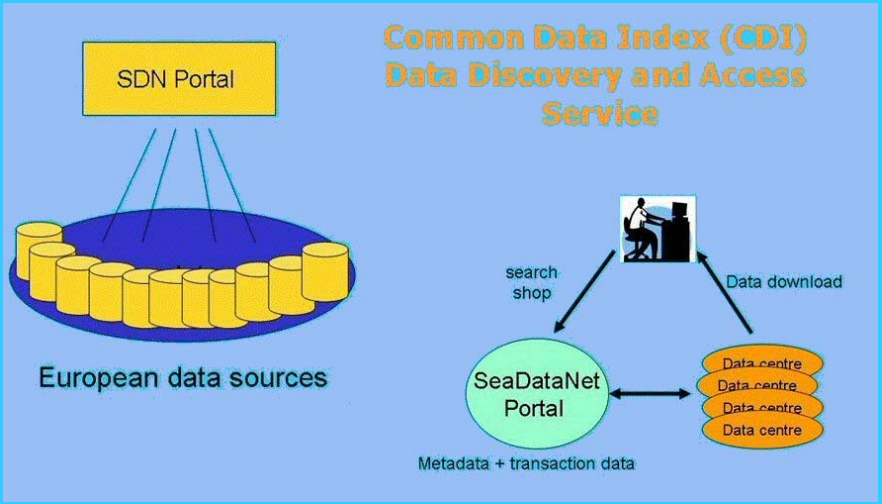
Option to combine multiple search criteria and free text search

Quick Search

Drill down facet search – intuitive and more easy for 1st time users

Data Access Service

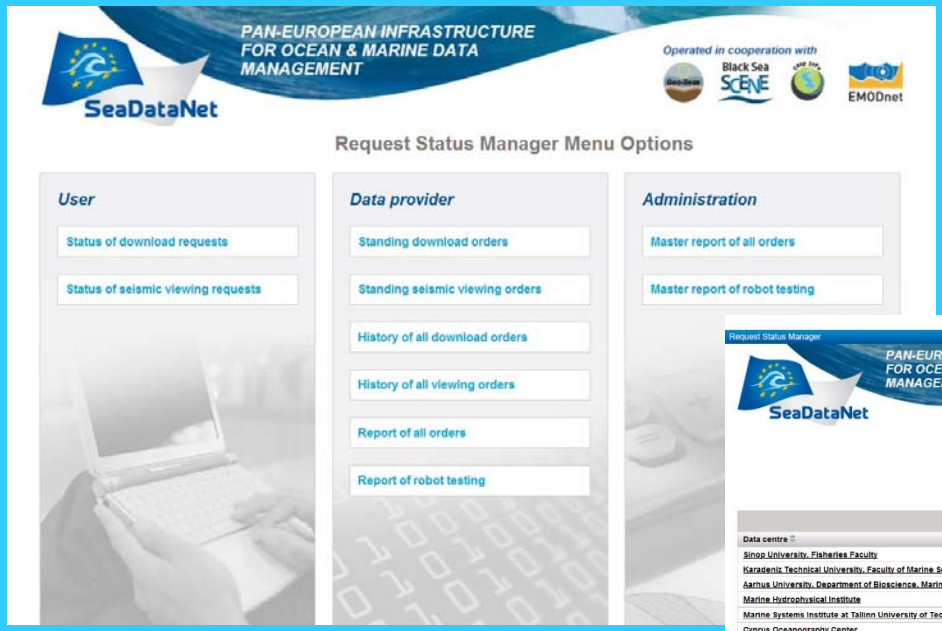
- An intelligent middle tier connection is configured between the SeaDataNet portal and the local data management systems at each of the data centres
- A shopping basket (max 10.000 records per basket) allows users to submit a shopping request for multiple data providers in one go and to follow its processing by each of the providers *via* an online transaction register.



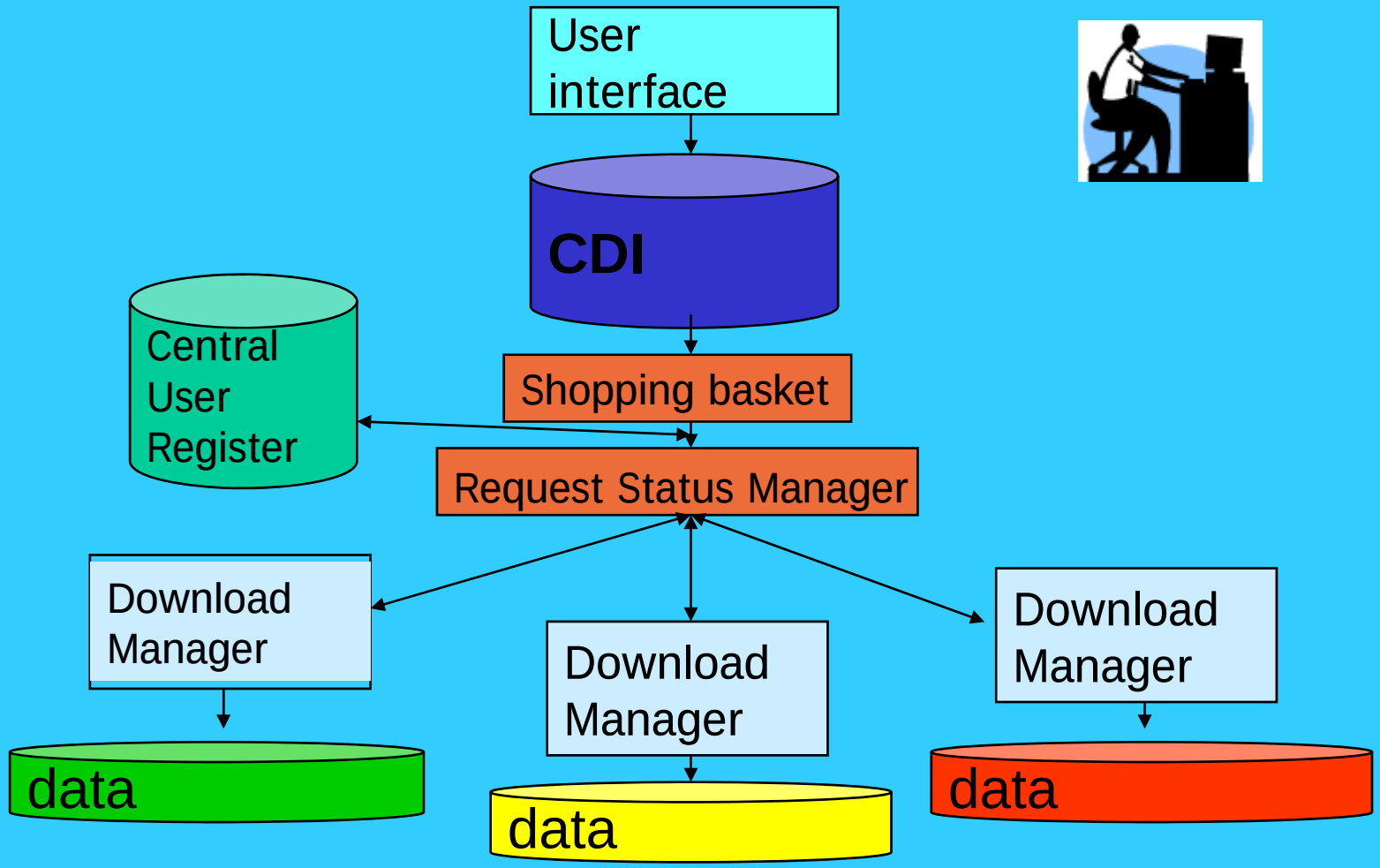
Data Access – Shopping dialogue

Request Status Manager (RSM)

- Tracking and tracing of all shopping requests by users, data providers and system managers
- Analysis of transactions and report generator
- Checking status of orders and downloading from data providers



CDI system components



Upgraded CDI service

User interface

CDI

Central User Register

Shopping basket

Services for quality management, versioning, transformation

Request Status Manager

Data cloud

Replication /download Manager

Replication /download Manager

Replication /download Manager

data

data

data

Upgrading CDI service using the cloud and HPC

- A cloud environment with High Performance Computing (HPC) facilities to host **copies of all data resources**
- Exchange by dynamic **replication** from the individual data centres, following their updating of the CDI catalogue service
- In the cloud buffer:
  - checking possible duplicates
  - Checking overall quality of formats
  - Checking integrity of data files and metadata relations.
  - Results of checks to be reported back to data centres for amendments of their submissions and/or local configurations for mapping data and metadata.
- Include versioning of data files
- Include transformation services for harmonizing data sets to common parameters and units, and converting data sets to other required output formats such as SeaDataNet NetCDF and relevant INSPIRE data models.

Upgrading CDI service using the cloud and HPC

- **Task WP9.1**: **Upgrading the CDI service making use of the cloud** – led by MARIS, EUDAT and IFREMER:
  - WP9.1.1: Specification of the SeaDataNet European cloud environment (EUDAT) – M6
  - WP9.1.2: Enhancing the EUDAT B2SAFE replication software (EUDAT, MARIS and IFREMER) – M12
  - WP9.1.3: Integrating and adapting the EUDAT B2HOST service (EUDAT and MARIS) – M12
  - WP9.1.4: Configuring the upgraded CDI service (MARIS and EUDAT) – M18
  - WP9.1.5: Deploying, testing and taking into operation the upgraded CDI service (MARIS and EUDAT) – M18
- **Task WP9.2: Developing and deploying additional services in the cloud for ensuring integrity and conformity of the CDI related cloud data resources -** led by IFREMER, MARIS, AWI and EUDAT - M18
- Upgraded CDI service fully deployed – M24 (*via* WP5 and WP3)

www.seadatanet.org