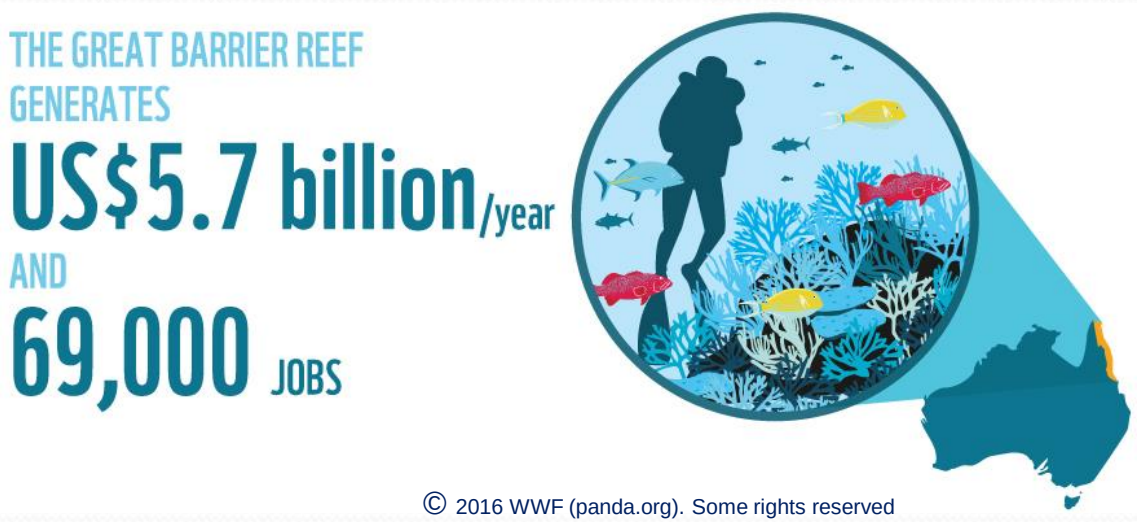
Ocean Data Interoperability Platform: ODIP

addressing key challenges for marine data management on a global scale

Helen Glaves, Dick M. A. Schaap, Jay Pearlman

Scientific and economic drivers for sharing of marine data

- Improved understanding of marine ecosystems
- Assessment of health of marine environments
- Modelling and forecasting of potential future changes
- Sustainable exploitation of the oceans
  - Maximization of ecosystem services
  - Minimization of human impact



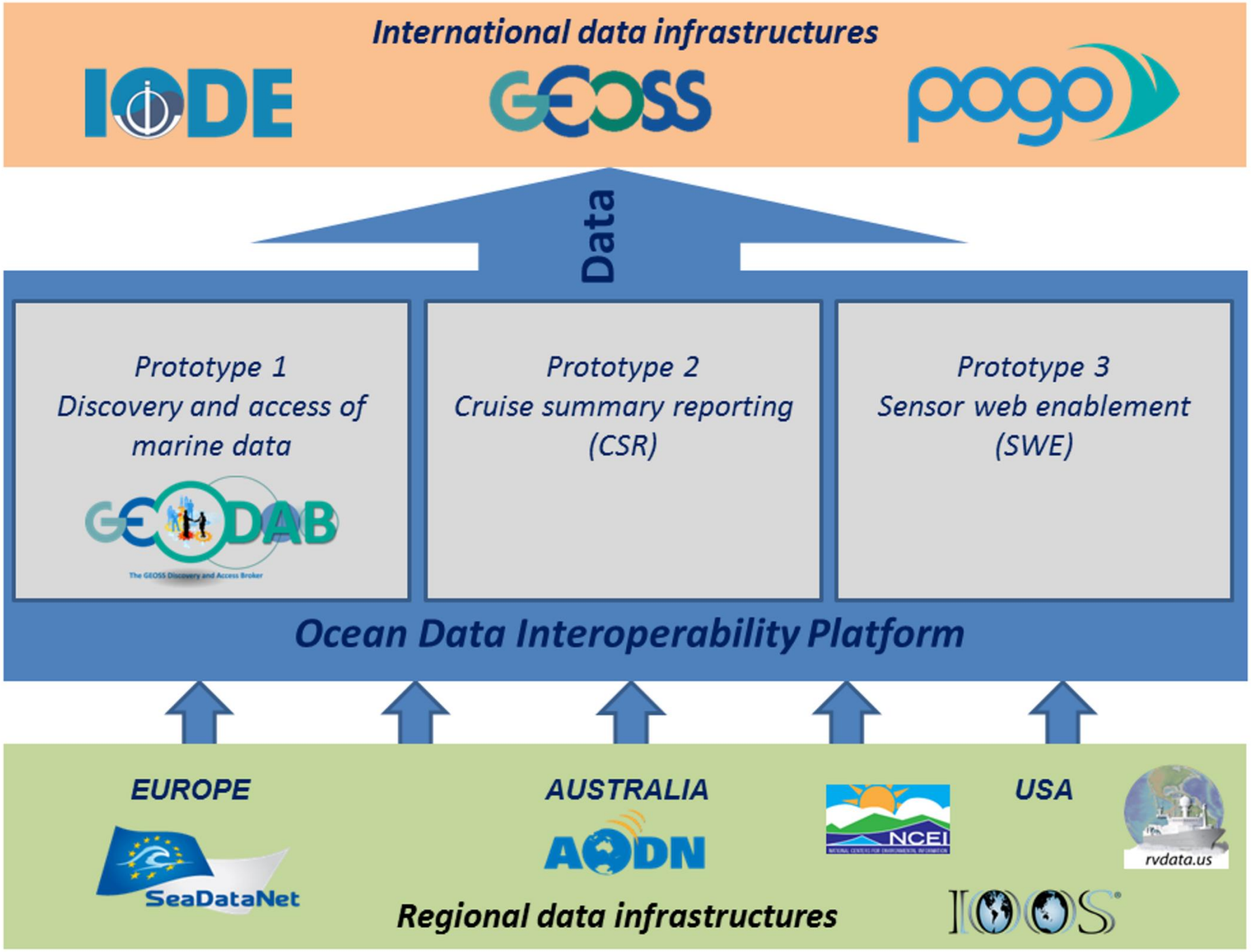
© 2016 WWF (panda.org). Some rights reserved

Marine e-infrastructures

- Regional e-infrastructures
  - Address specific ‘local’ requirements for data discovery and access
  - Are developed in response to needs of the user community and funding agency policy and guidelines
  - Created in isolation to those in other regions
- Global e-infrastructures
  - Domain specific e.g. IODE – Ocean Data Portal (IOC)
  - Multidisciplinary e.g. GEOSS

The challenges of use and re-use of marine data:

- Inconsistent data formats
- Spatial referencing
  - Coordinate systems
  - Horizontal datums
- Different standards/best practice
- Data access policies (organisational, national and regional level)


International data infrastructures
IODE
GEOSS
pogo
Data
Prototype 1
Discovery and access of marine data
GEODAB
The GEOSS Discovery and Access Broker
Prototype 2
Cruise summary reporting (CSR)
Prototype 3
Sensor web enablement (SWE)
Ocean Data Interoperability Platform
EUROPE
SeaDataNet
AUSTRALIA
AODN
NCEI
USA
rvdata.us
IOOS
Regional data infrastructures

Project Objectives: ODIPII

- Continue and extend the activities of the previous ODIP project
- Development of joint prototype activities
  - further development of the existing prototypes to fully operational systems
  - additional prototype development tasks
- Extend scope of the project to include:
  - other domains e.g. biological oceanography
  - regions e.g. Canada
  - additional partners

List of ODIP II topics

- Continuation and extension of the three current ODIP prototype development tasks.
  - Prototype 1: Interoperability of regional discovery services towards GEOSS and ODP: improving and including marine biology;
  - Prototype 2: Cruise Summary Reporting towards POGO: expanding scope and coverage;
  - Prototype 3: Sensor Web Enablement (SWE): making further progress towards common marine profiles and services and interaction with industry;
- Continuation and extension of the current cross-cutting topics:
  - Controlled vocabularies: content, RDF, SPARQL endpoints, user interfaces, mappings, ..
  - Data publishing, data citation, interaction with data management
  - Persistent identifiers for data and persons

Discovery and access of marine data

- Integration of regional data discovery and access services
- Utilizing GEO-DAB broker
- Flexible solution
- High degree of local autonomy
- Single point of access for data discovery
- Metadata exposed in global systems: GEOSS and ODP

Status of Discovery and Access

- Established interoperability between SeaDataNet CDI, US NODC (NCEI), and IMOS Marine Community Profile (MCI) data discovery and access services using a brokerage service, and moving towards interacting with the global IODE-ODP and GEOSS portals, led by SeaDataNet (Europe);
- Initial steps were for metadata interoperability supporting discovery. Current activities in ODIP II are addressing both data access and discovery.
- Focus now is on semantic interoperability

Cruise summary reports

- Report on cruise or field experiments at sea
- Regional cruise summary reporting systems
- Global cruise catalogue (POGO)
- Adoption of common ISO19139 compliant schema
- Supports cruise planning

R.R.S. 'Boaty McBoatface'

Cruise summary reports

- Establishing a common deployment and interoperability between Cruise Summary Reporting systems in Europe, US and Australia by making use of GeoNetWork, and moving towards interaction with the global POGO portal (http://www.pogo-oceancruises.org)

- A set of 130 CSRs for the USA research vessels *Kilo Moana* and *Falkor* were harvested and successfully integrated into the POGO CSR Catalog service.

Sensor web enablement

- (Near) real-time access to ocean sensor data
- OGC suite of standards and best practice
    - Domain independent
    - High degree of flexibility
    - Potential for divergence
- Common marine SWE profile
- Delivers ocean observing system data directly to data centres
- Supports integration with data discovery systems

Sensor web enablement Implementation

- Establishing a prototype for a Sensor Observation Service (SOS) and formulating common O&M and SensorML profiles for selected sensors (SWE), installed at research vessels and in real-time monitoring systems.
- In ODIP II, OGC SWE standards will be evaluated for the interoperable sharing of oceanographic observation data and metadata. This includes also attention for use of appropriate controlled vocabularies for marking up the metadata profiles in a consistent way.

Other possible ODIP II topics - 1

- netCDF-CF standardization and interoperability with OGC data models and protocols
- netCDF4 and openDAP support for the SeaDataNet gridding tool DIVA
- Linked data principle and applications
- Harvesting from several data sources, automated aggregation (duplicates elimination), quality control, harmonisation, gridding and visualisation.
- Integration of data services with modeling workflows

Other possible ODIP II topics - 2

- Automated processing of near real time data streams, particularly through the use of Web Processing Services (WPS) and workflow tools;
- Big data – Web Processing Services - Cloud systems;
- Data ingestion systems;
- Common standards for capture of provenance information;
- Real-time observation systems

Final remarks - 1

- Ecosystem level marine research has made an integrated global network of data services a necessity

- Leveraging development activities of existing marine e-infrastructures to establish a common global frame-work for marine data management potentially overcomes many of the recognised barriers to the sharing of marine data.

Final remarks - 2

- An approach based on existing data systems is one that is potentially highly scalable and robust, as well as being transferable to other domains
- Establishing interoperability across regional data infrastructures and with the larger global data systems makes marine data more widely available for a diverse range of multidisciplinary applications