eCUDO.pl

Oceanographic Data and Information System for Polish NODC

Marcin Wichorowski	Institute of Oceanology Polish Academy of Sciences
Mirosława Ostrowska	Institute of Oceanology, Polish Academy of Sciences
Michał Piotrowski	Maritime Institute in Gdańsk
Lena Szymanek	National Marine Fisheries Research Institute
Urszula Pączek	Polish Geological Institute National Research Institute
Michał Wójcik	Maritime Institute in Gdańsk

CODATA: Towards Next-Generation Data Driven Science 2019

Beijing, September 19th-20th, 2019

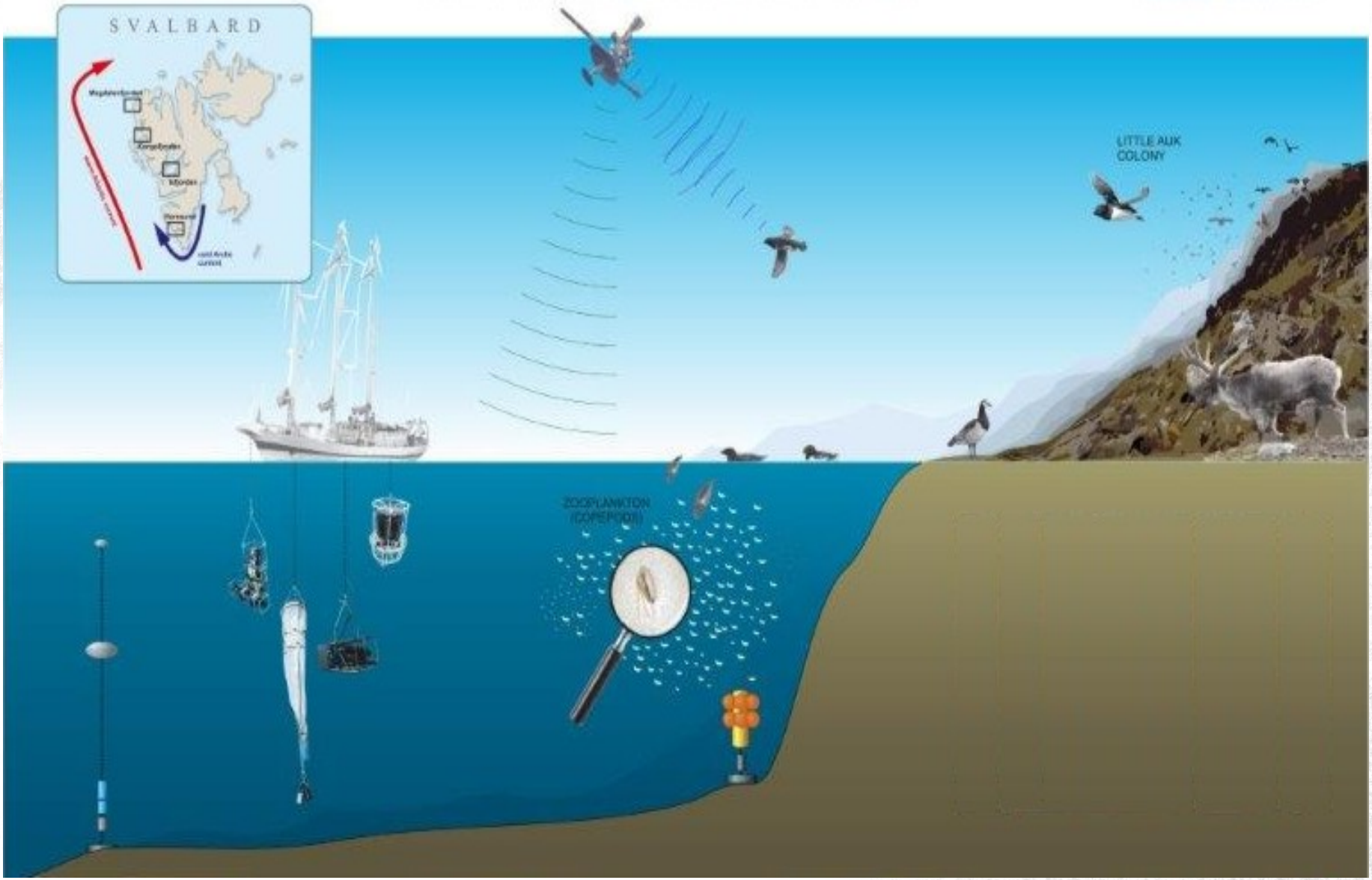
Fundusze Europejskie Polska Cyfrowa

Rzeczpospolita Polska

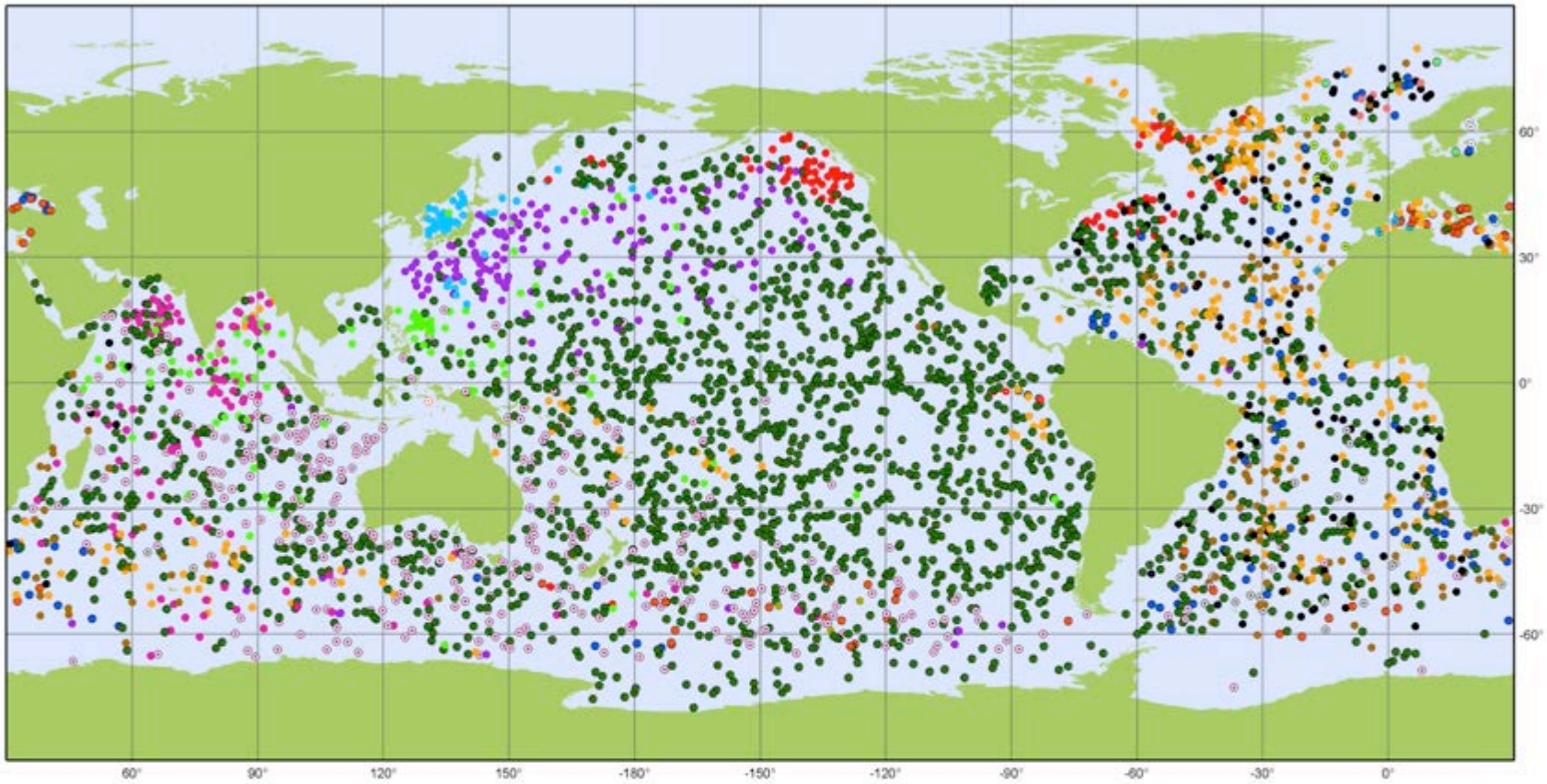
eCUDO.pl

Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Complexity of oceanographic research



Complexity of oceanographic research



Argo

National contributions - 3881 Operational Floats

Latest location of operational floats (data distributed within the last 30 days)

February 2018

ARGENTINA (1)	EUROPE (94)	INDIA (124)	KENYA (1)	PERU (3)	USA (2179)
AUSTRALIA (361)	FINLAND (3)	INDONESIA (1)	MEXICO (2)	POLAND (5)	
BRAZIL (3)	FRANCE (277)	IRELAND (12)	NETHERLANDS (24)	KOREA, REPUBLIC OF (53)	
CANADA (87)	GERMANY (142)	ITALY (65)	NEW ZEALAND (6)	SPAIN (5)	
CHINA (105)	GREECE (2)	JAPAN (156)	NORWAY (7)	UK (163)	

Generated by www.jcommops.org, 02/03/2018

Project partners

Maritime Institute in Gdańsk

Polish Geological Institute
National Research Institute

Sea Fisheries Institute
National Research Institute

Institute of Oceanology
Polish Academy of Sciences

University of Gdańsk

Pomeranian Academy in Słupsk

University of Szczecin

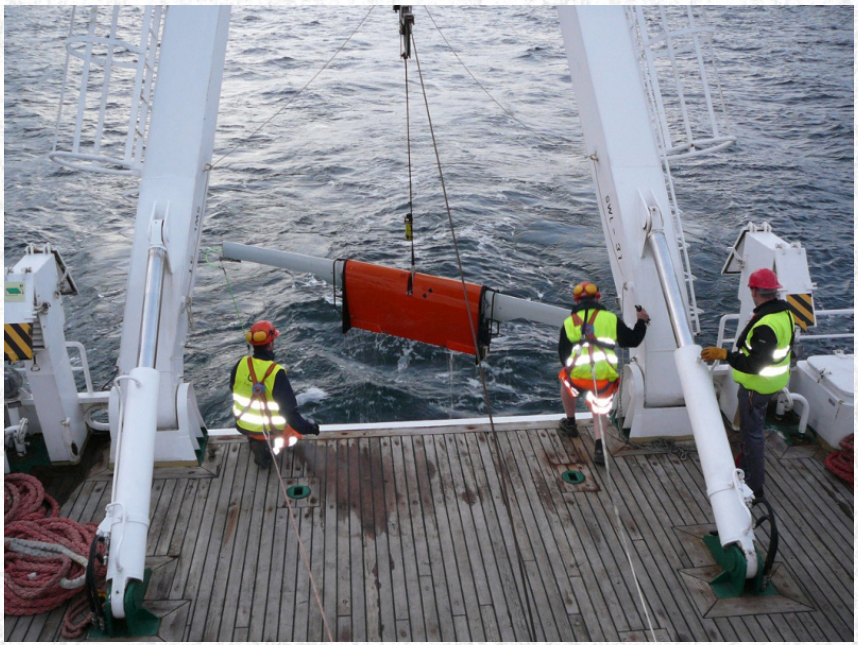
Baltic Bottom Base

- BalticBottomBase is an Integrated Information Platform on South Baltic Environment which was launched in 2014 by Maritime Institute in Gdańsk
- It is an on-line database which makes available some data gathered by Maritime Institute in Gdansk
- Most of the available data are oceanographic model data – more than 50 TB of data
- It has more than 150 registered users
- As time goes by, it is harder to maintain the system as IT environment is constantly changing

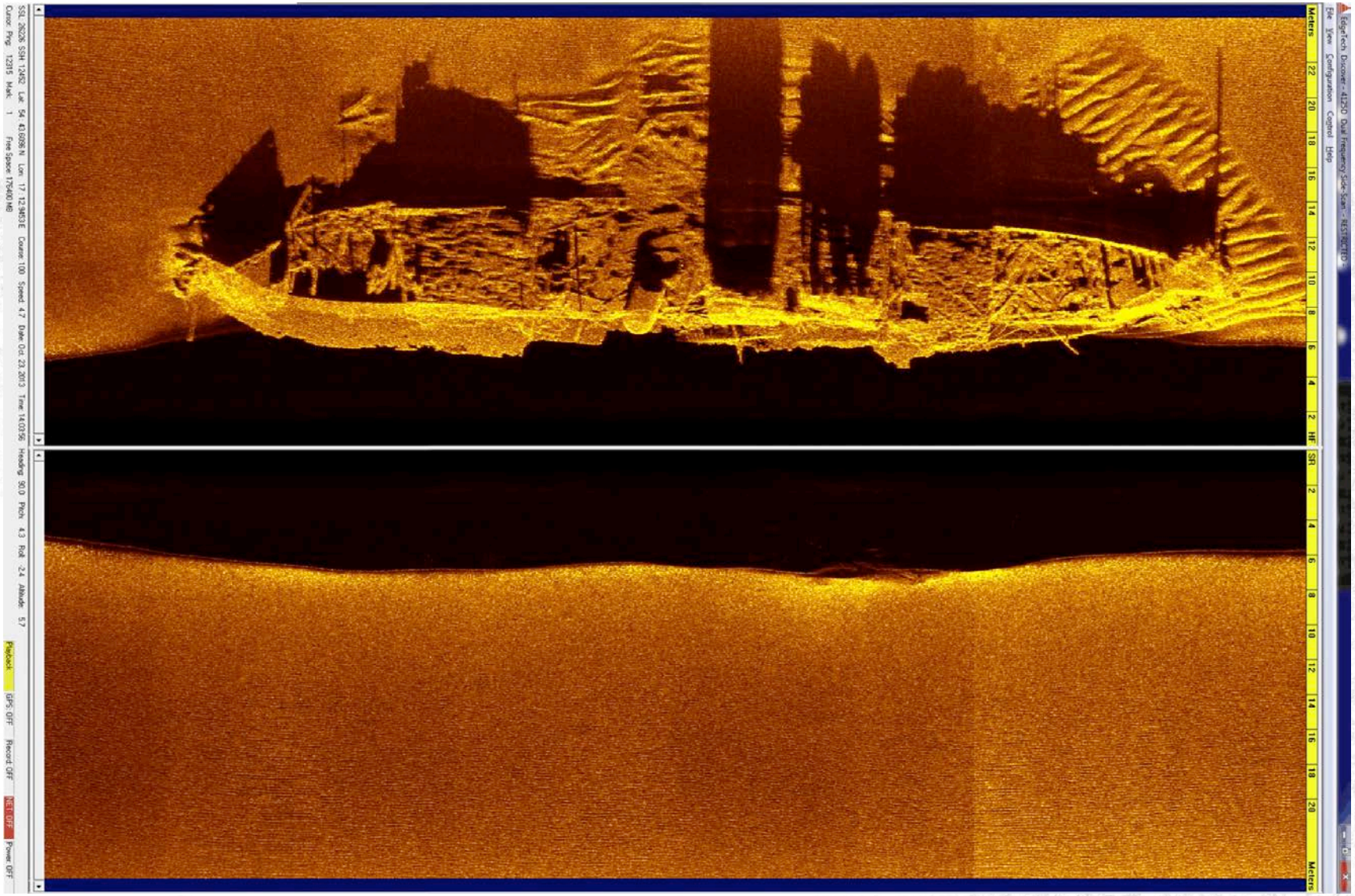
BalticBottomBase

Types of data acquired by Maritime Institute in Gdańsk

- single beam echo sounder profiles
- multi beam echo sounder profiles
- side scan scanner profiles
- sub-bottom profiler data
- magnetometer data
- oceanographic model data

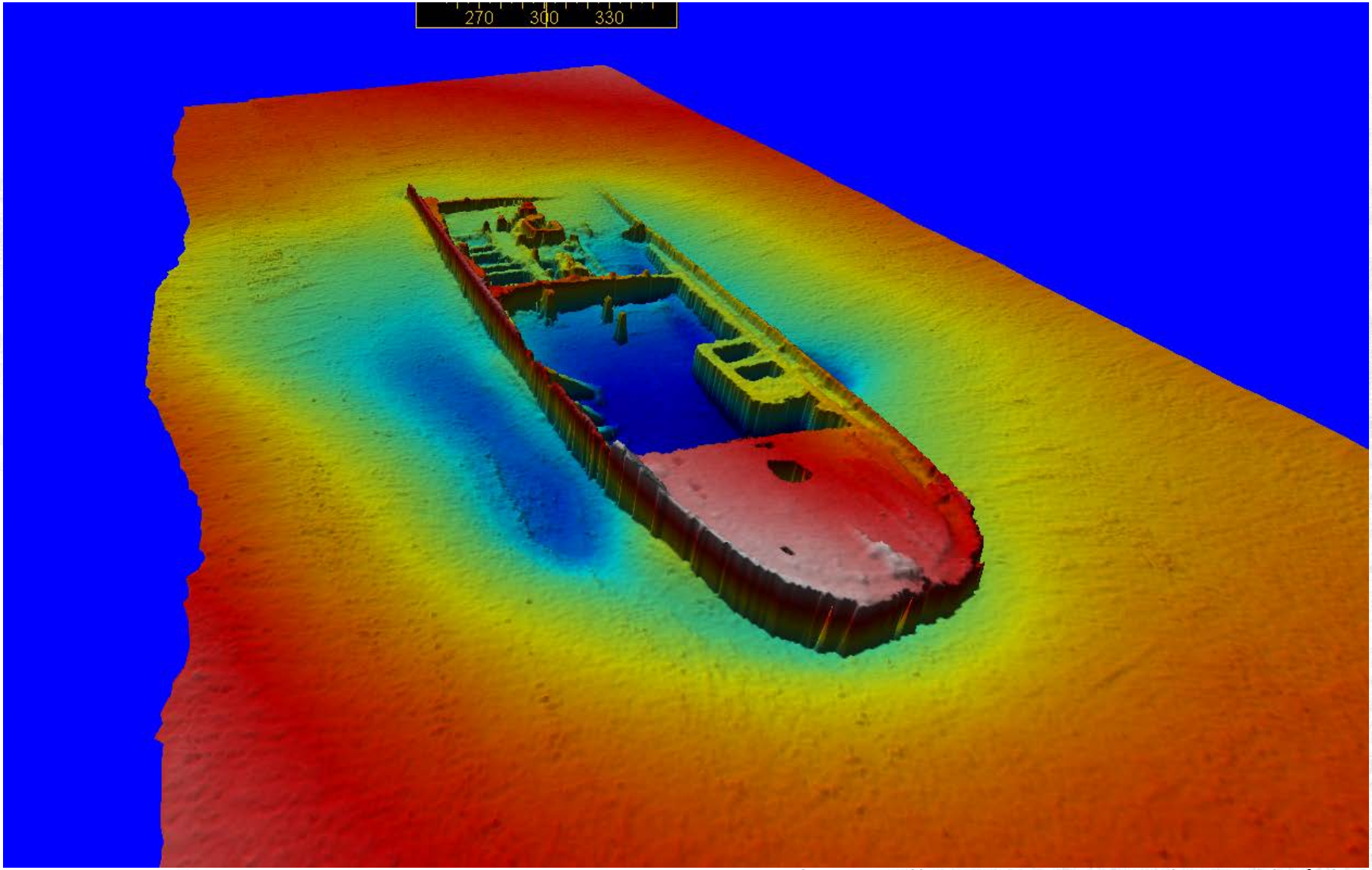


Types of data acquired by Maritime Institute in Gdańsk



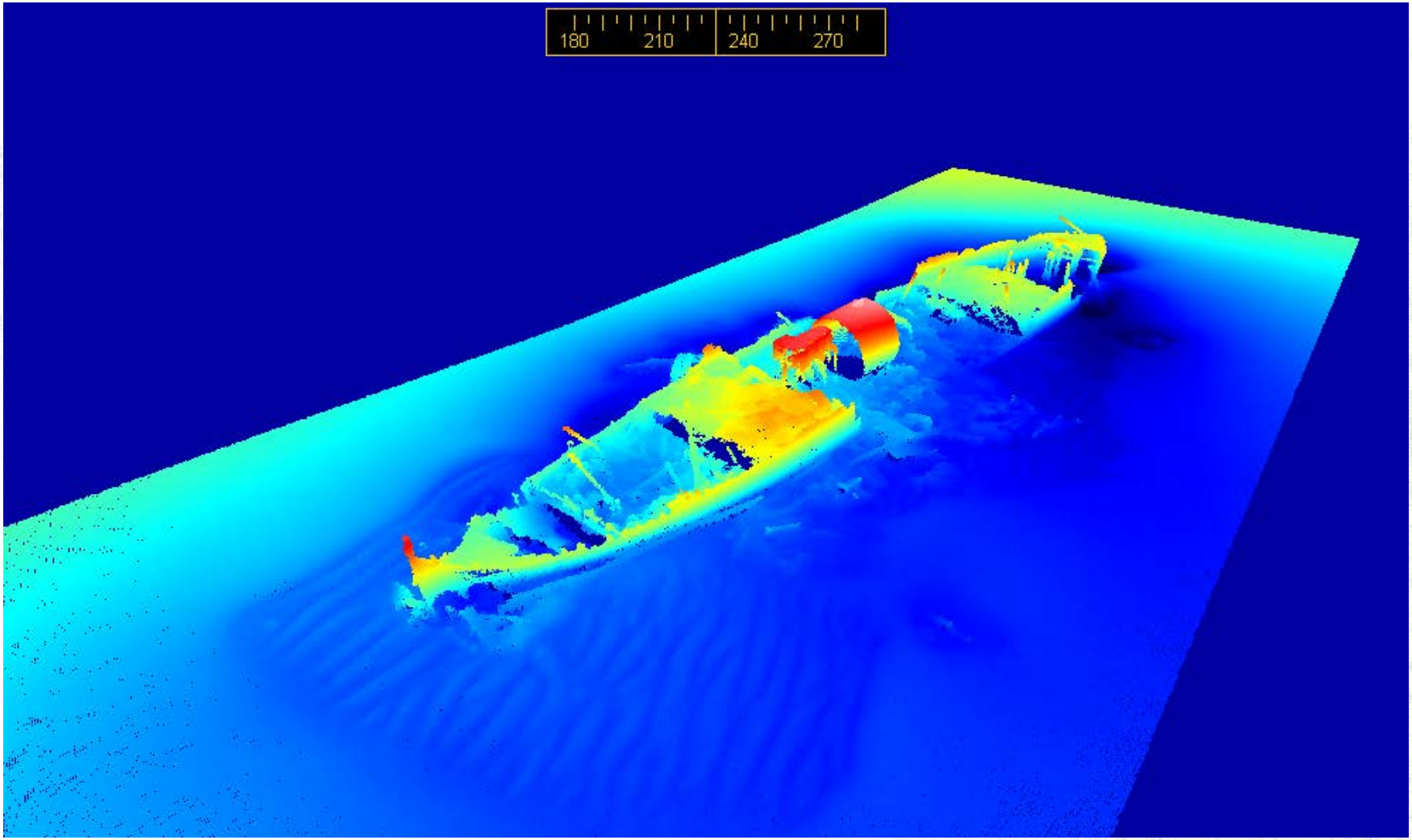
Autor: Jacek Koszałka, Instytut Morski w Gdański

Types of data acquired by Maritime Institute in Gdańsk



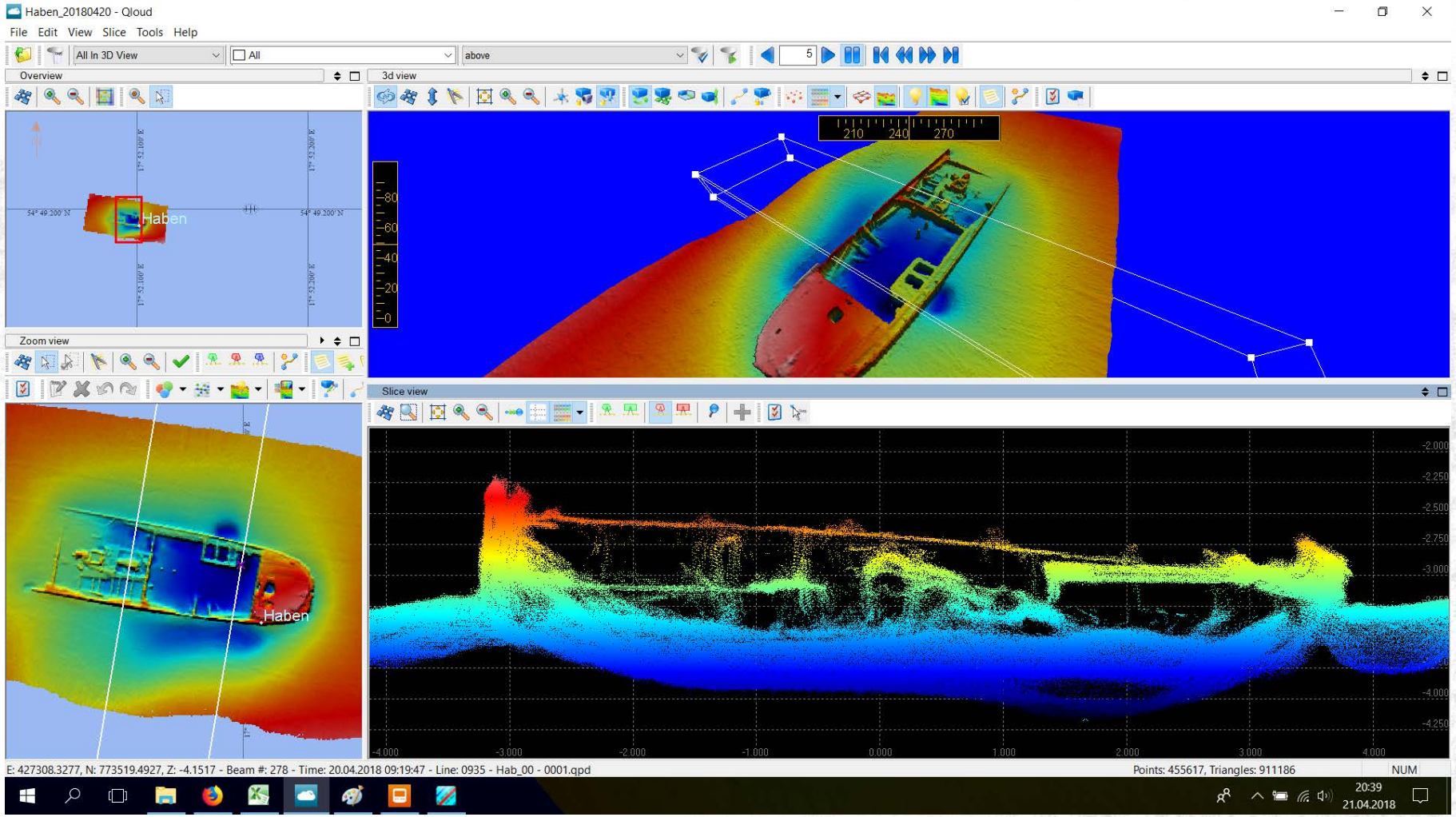
Autor: Jacek Koszałka, Instytut Morski w Gdański

Types of data acquired by Maritime Institute in Gdańsk



Autor: Jacek Koszałka, Instytut Morski w Gdański

Types of data acquired by Maritime Institute in Gdańsk



Autor: Jacek Koszałka, Instytut Morski w Gdański

Data are waiting to be open

More than:

- 500 SBES profiles
- 1700 MBES profiles
- 600 SSS profiles
- 400 SBP profiles
- 30 MAG profiles

are waiting to overcome legal and technical obstacles to be fully opened.

- Legal obstacles:
  - audit of legal agreements for projects
- Technical obstacles:
  - labour effort needed to provide metadata
  - systems for data storage and dissipation

Geological and geophysical database of Polish EEZ

- PGI-NRI joined **Geo-Seas** *Pan-European Infrastructure for marine and ocean geological and geophysical data* project (2009-2013)
- One of 26 marine geological and geophysical data centres
- Metadata compliant with INSPIRE, SeaDataNet dictionaries and ISO 19115 standards

EMODnet GEOLOGY
Discover Europe's marine geology

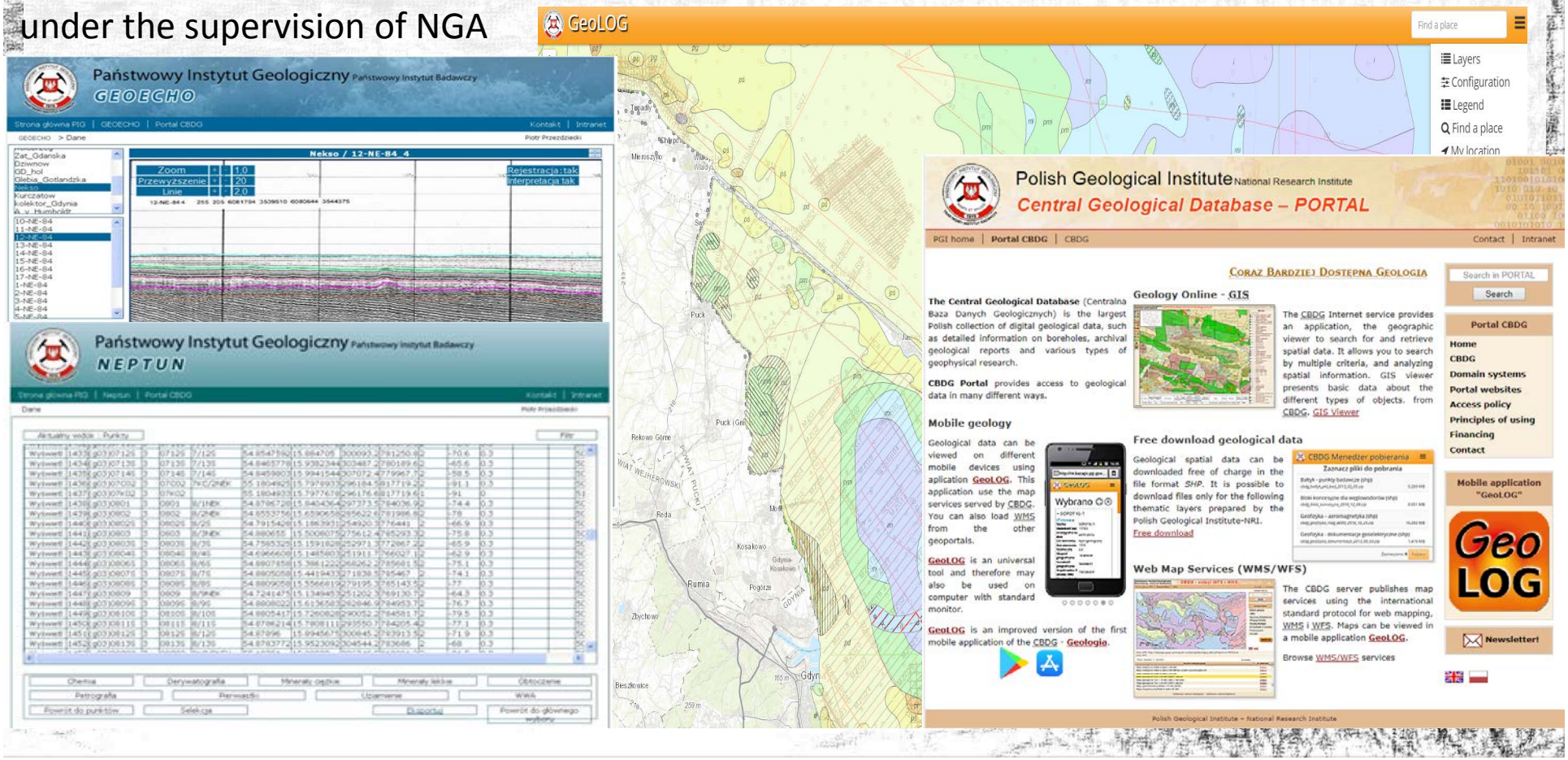
- PGI-NRI joined **EMODnet** *European Marine Observation and Data Network* project (2009, still, 3rd edition), long-term, EU's supported
- Discipline-based theme – **EMODnet Geology**
- Bringing together harmonised offshore data (sea-floor geology, seabed substrates, rates of coastline migration, geological events and probabilities and mineral resources)
- Providing free access to geological data and metadata held by various organizations in Europe based on international standards and geological data products compiled at a scale of 1:250,000.

Geological and geophysical database of Polish EEZ

National Geological Archives (NGA): collects, maintains and provides geological, hydrogeological and geophysical materials, maps and drill cores.

Central Geological Database (CGD) provides access to digital geological data via its PORTAL: geology Online – GIS viewer, GeoLOG – mobile viewer, Web Map Service (WMS/WFS) and SHP

NEPTUN and GeoEcho – CGD PORTAL domains, GeoLOG and GIS viewer layers, wms, wfs under the supervision of NGA



Geological and geophysical database of Polish EEZ

NEPTUN – sampling points

- Points metadata: goal, geodesy, kind
- Samples metadata
- Analyses results
- Core macroscopic description

GEOECHO – geophysical data

- Profile metadata: goal, geodesy
- Equipment description
- Rasters of registration and interpretation

GEOMORZE – GIS DATABASE

- Bathymetric maps with DTM
- Geological maps
- Geochemical maps
- Grain size maps
- Etc.....

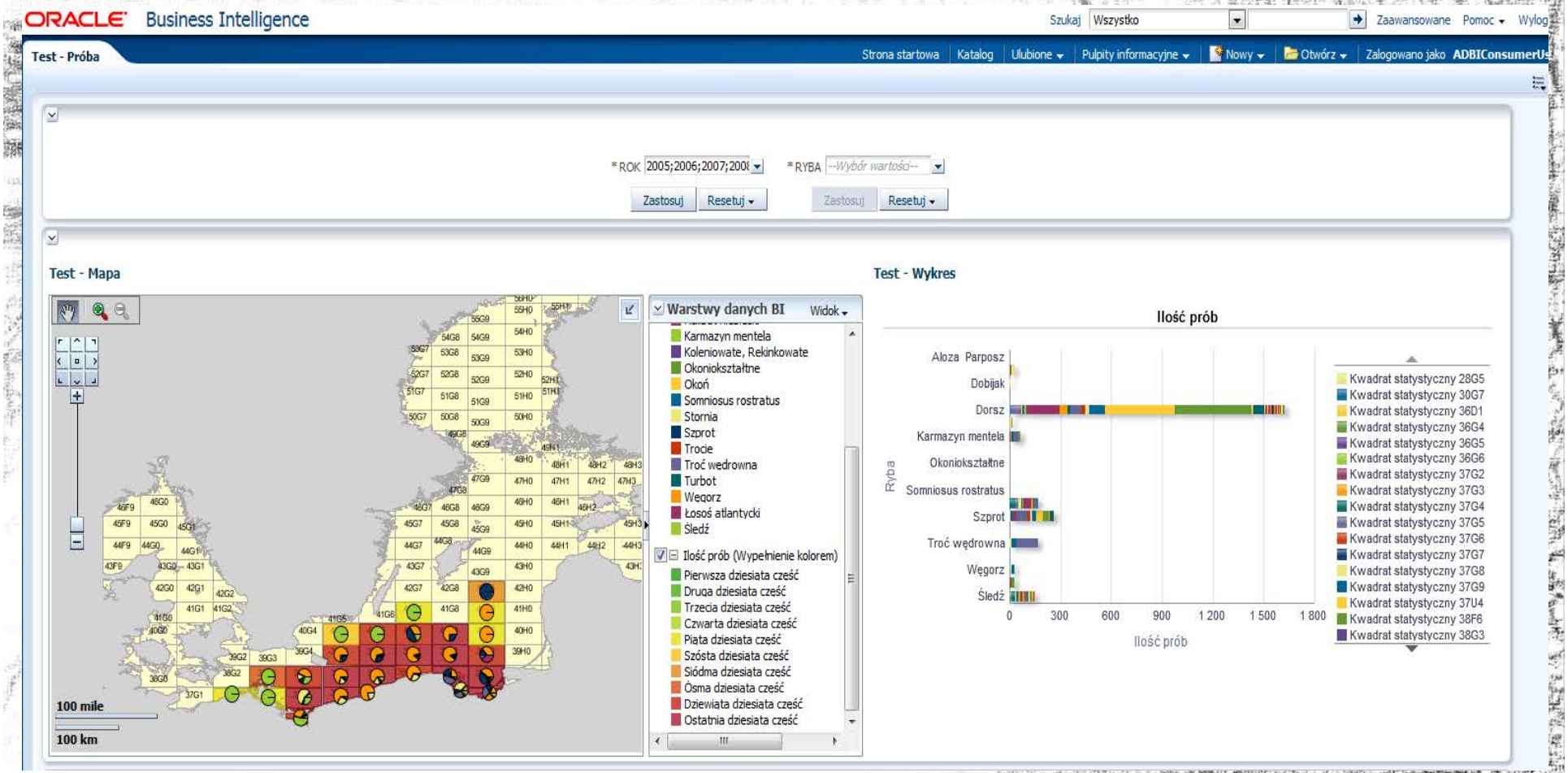
SOURCE DATA REPOSITORY

- Source clouds of point and processed data from terrestrial and airborne laser scanning
- Digital geophysical registration
- Sonar mosaics, bathymetry

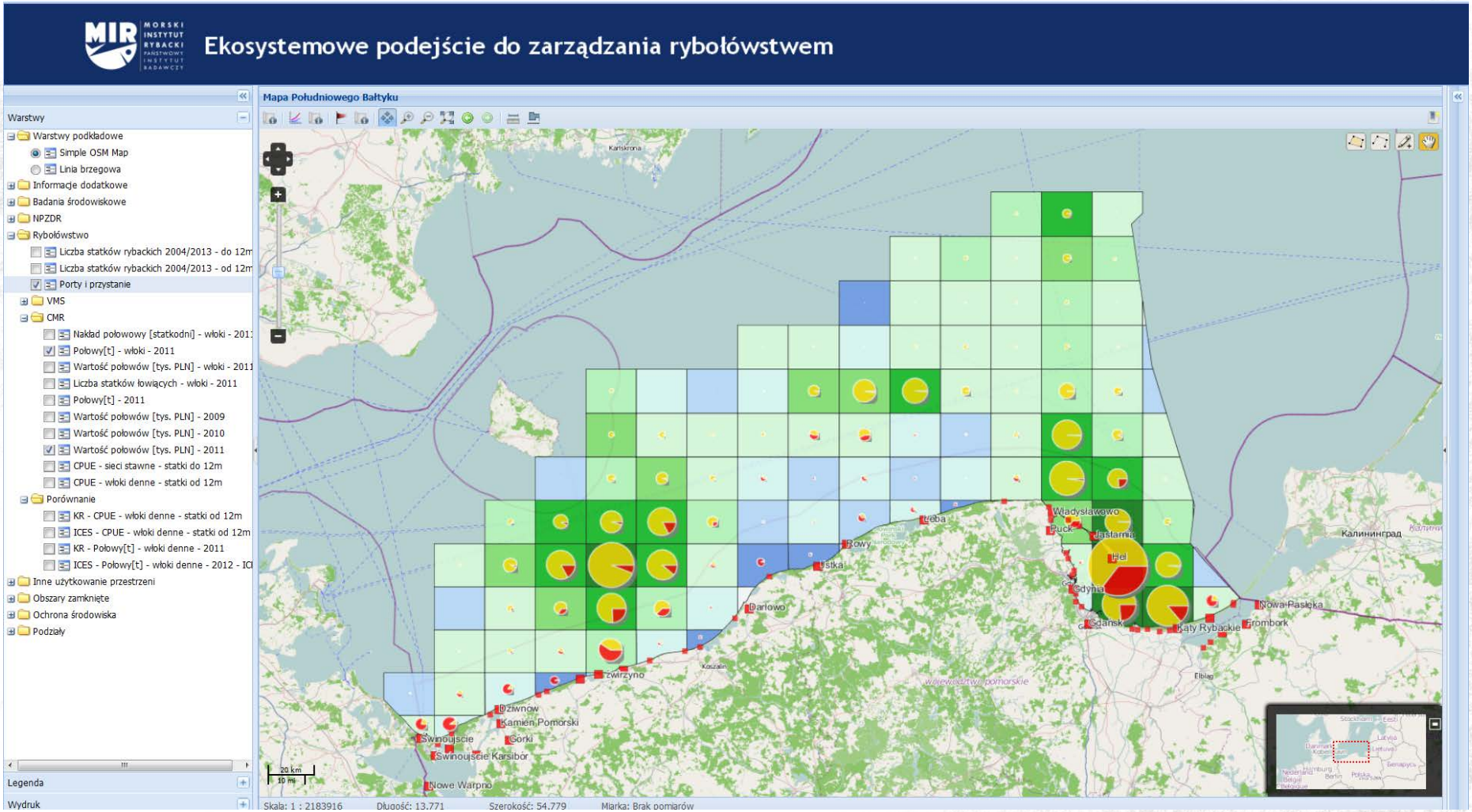
EXTERNAL DATA REPOSITORY

- Meteorological data (IMWM-NRI)
- Sea level (IMWM-NRI)
- Etc...

Sea Fisheries Institute – Scientific Data Repository

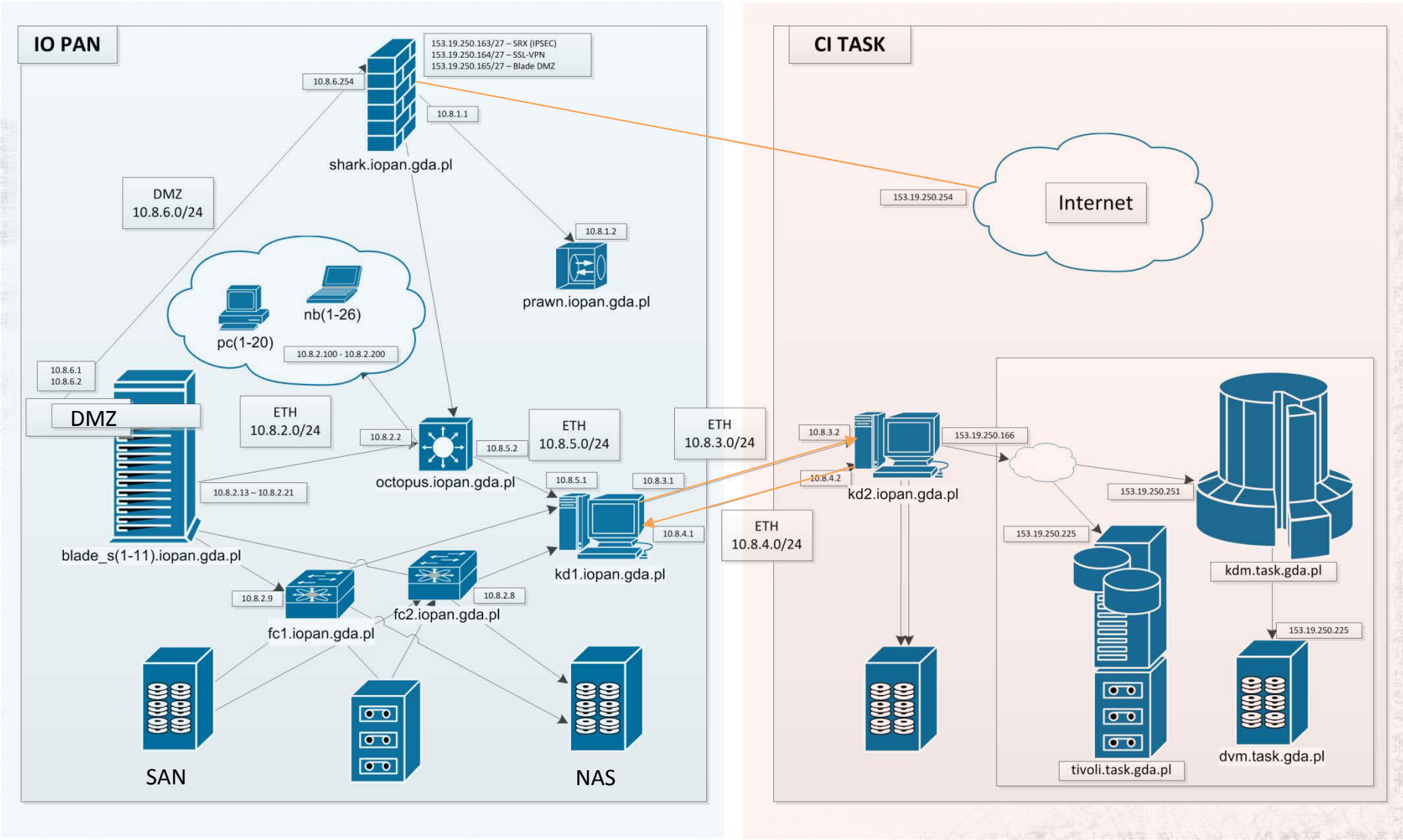


Sea Fisheries Institute – Scientific Data Repository



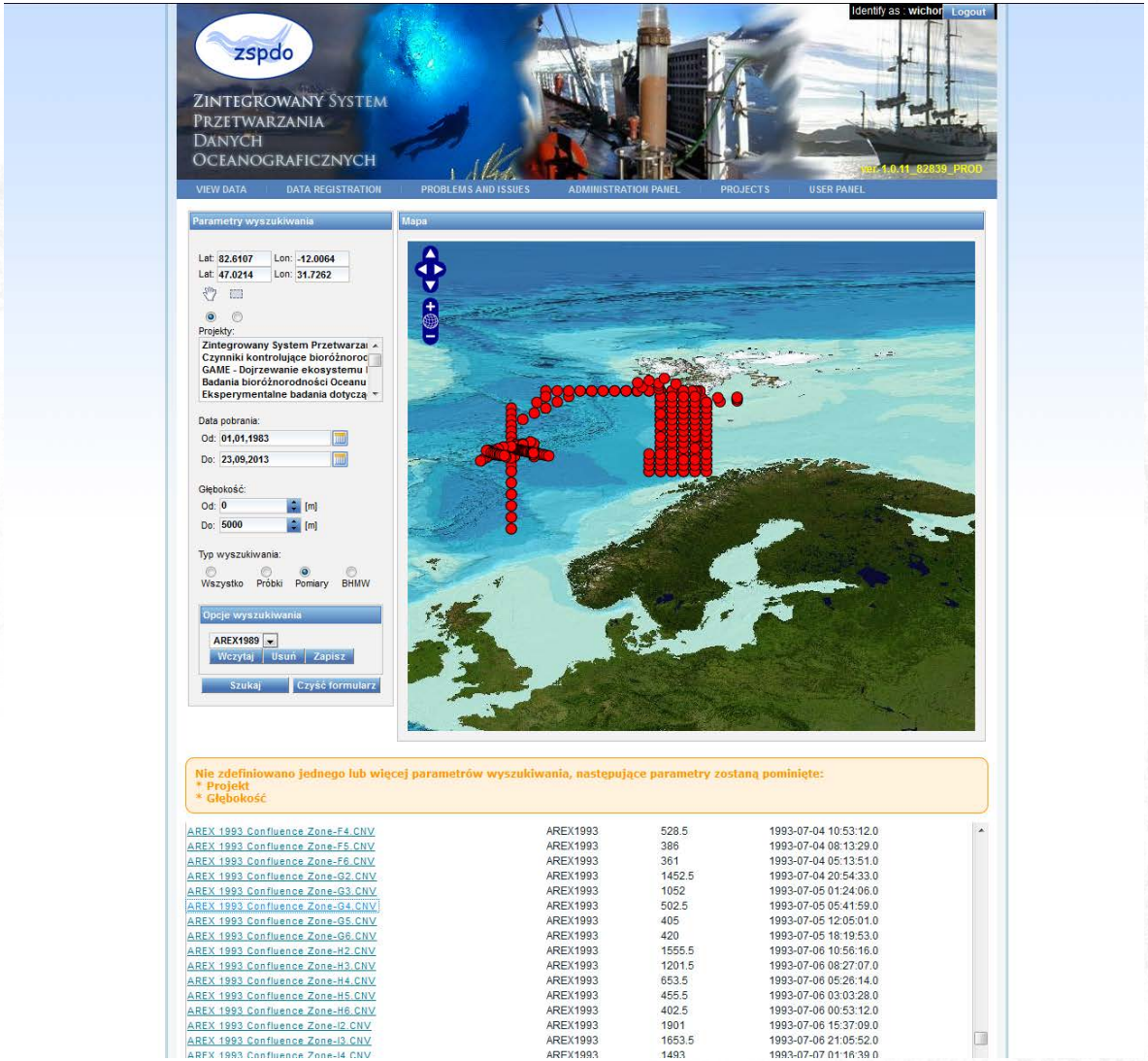
Data management infrastructure at IO PAN

Integrated Oceanographic Data and Information Management Sytem

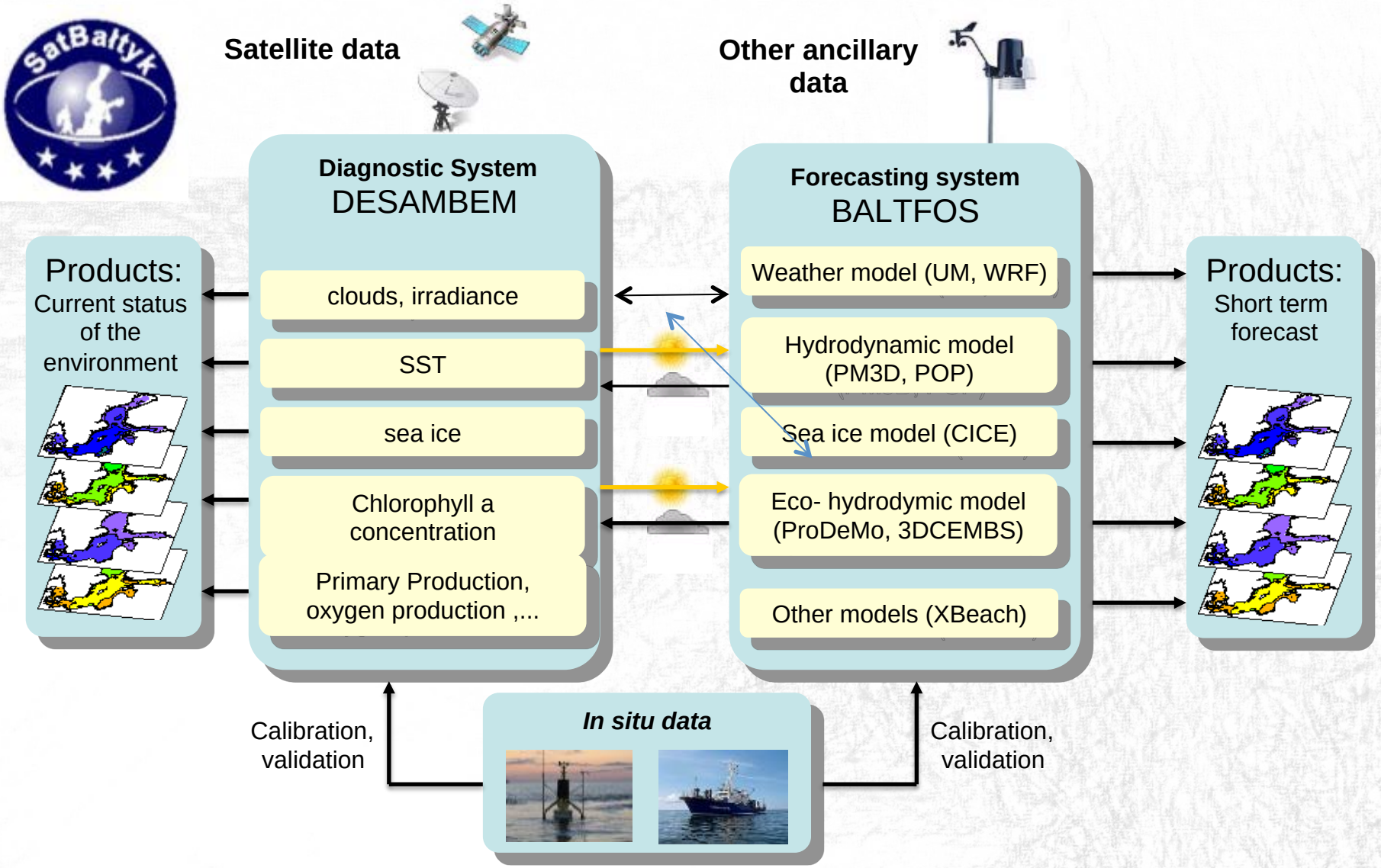


Data management infrastructure at IO PAN

Integrated Oceanographic Data and Information Management Sytem

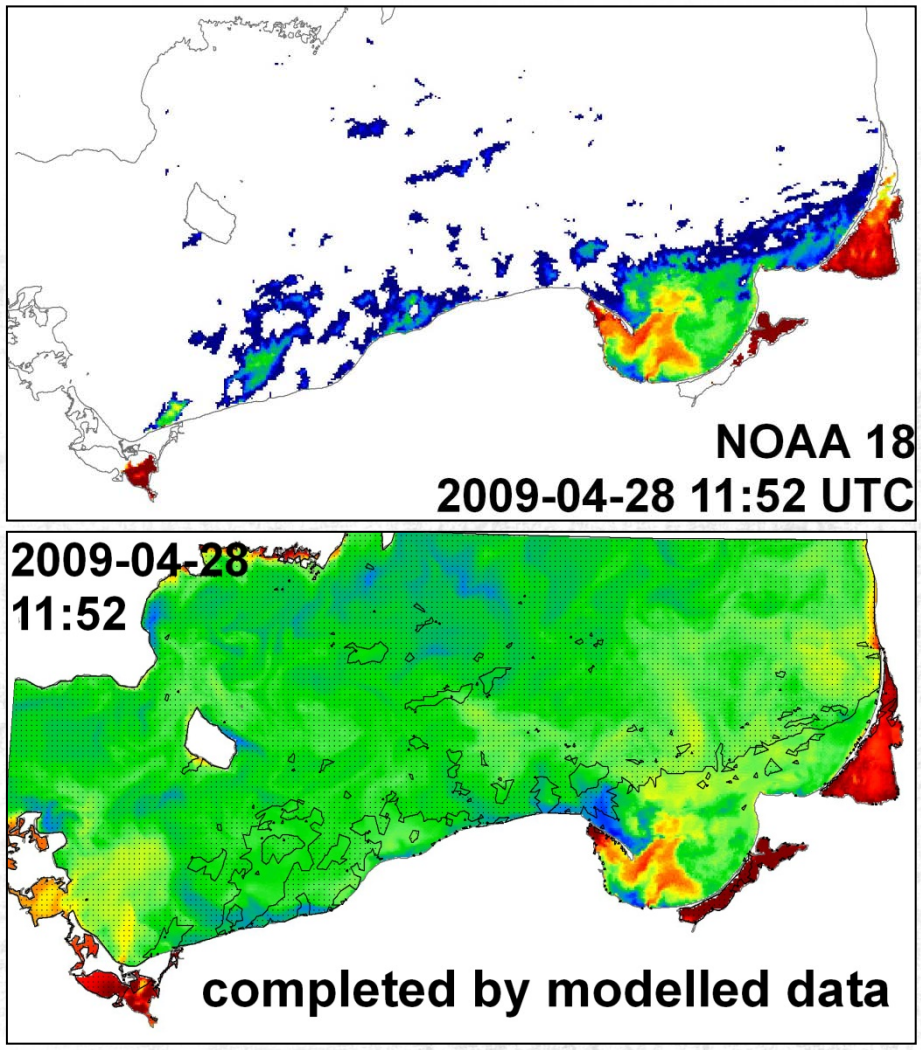


SatBaltyk system data flow

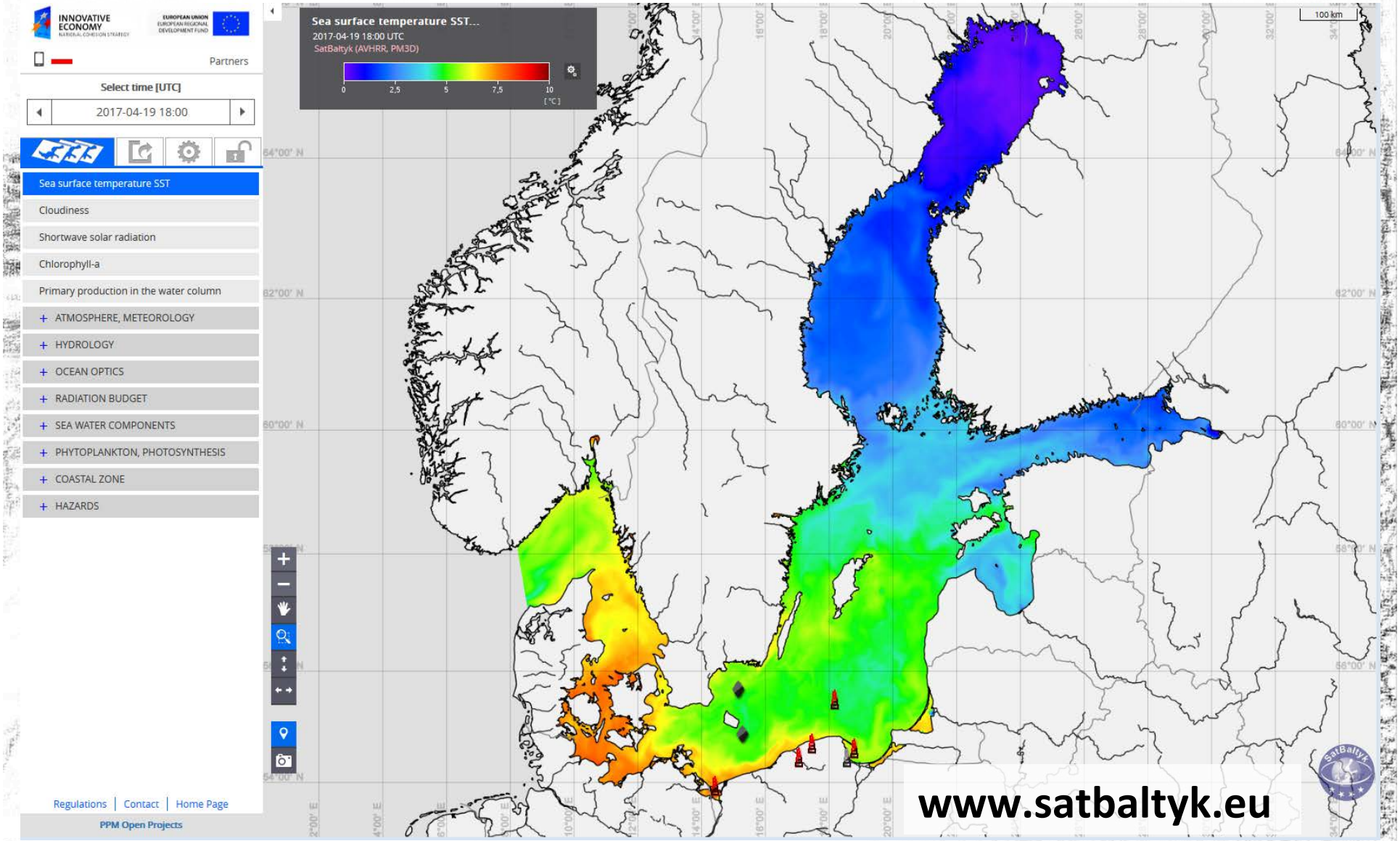


Merged Sea Surface Temperature Map

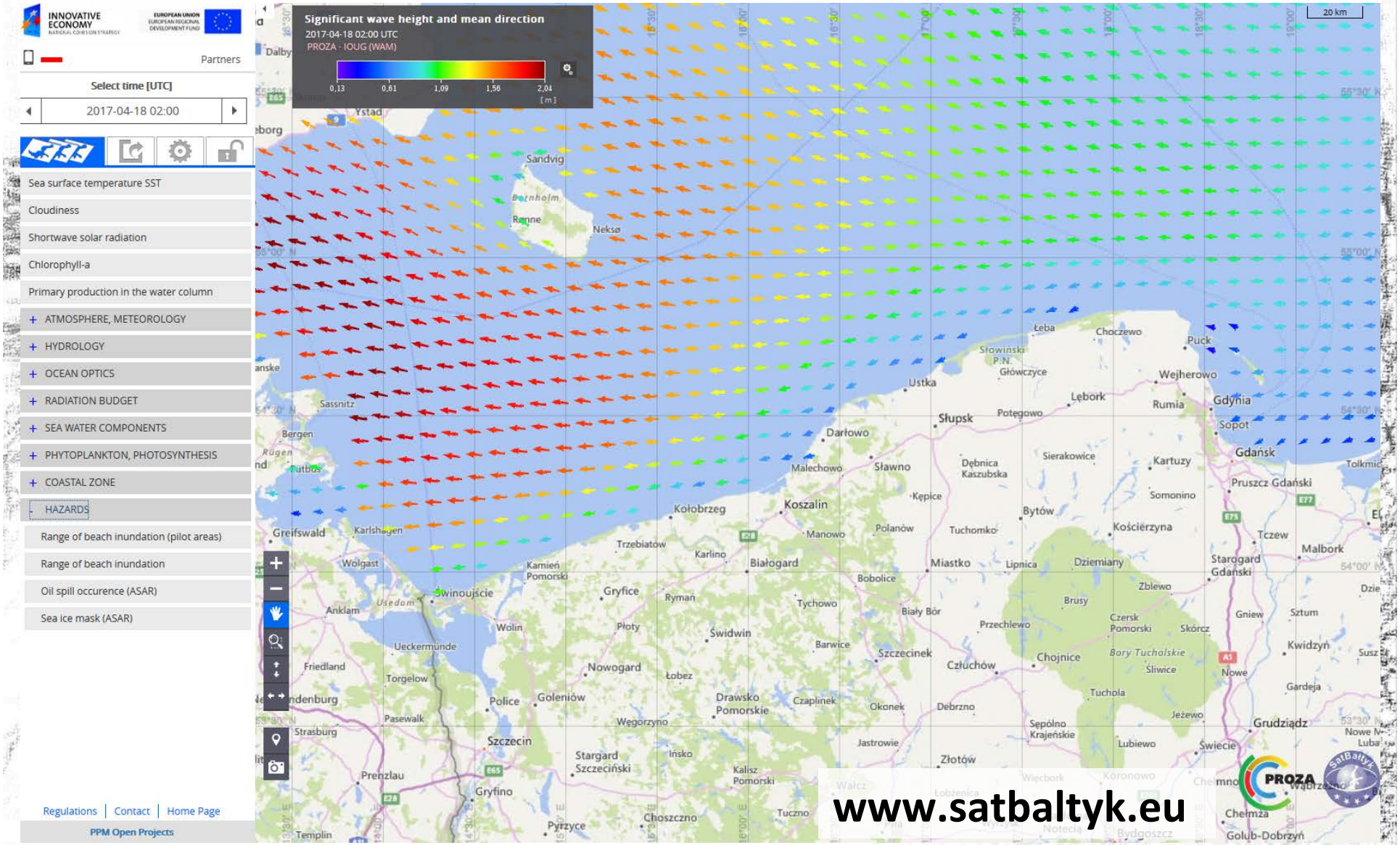
Sea surface temperature determined on the basis of remotely sensed data and the M3D hydrodynamic model (grid resolution 0.5 NM) when a large part of the sky over the sea is overcast



SatBaltyk product portal

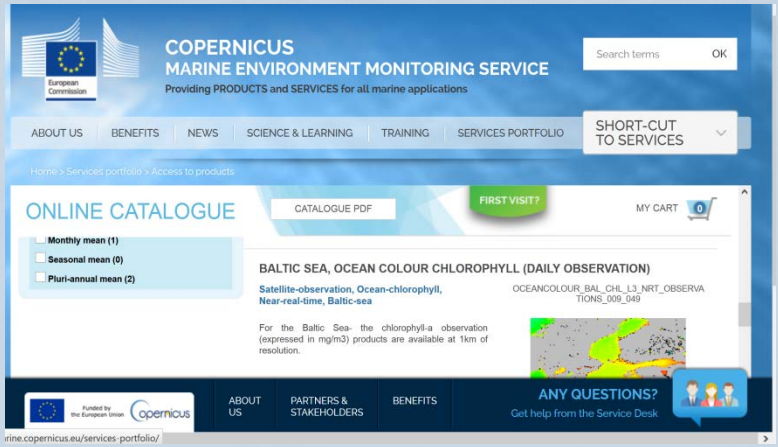


SatBaltyk system product portal



Complementarity of SatBaltyk and Copernicus services

Copernicus Marine Environment Monitoring Service

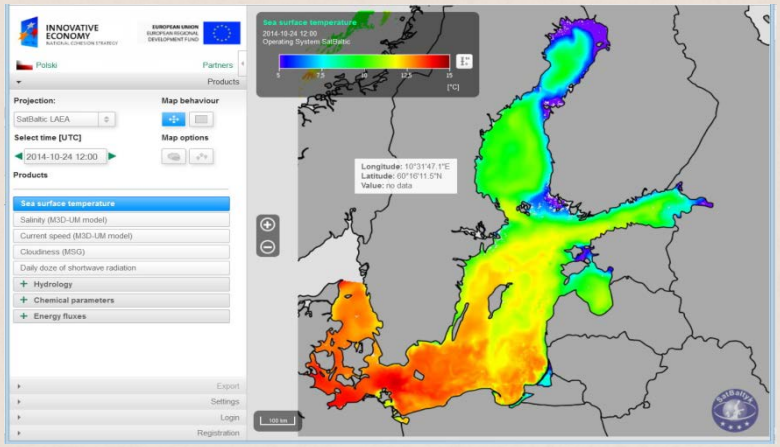


+ long time support by EC

+ all European marine waters

vs.

SatBaltyk: – Baltic Environmental Satellite Remote Sensing System



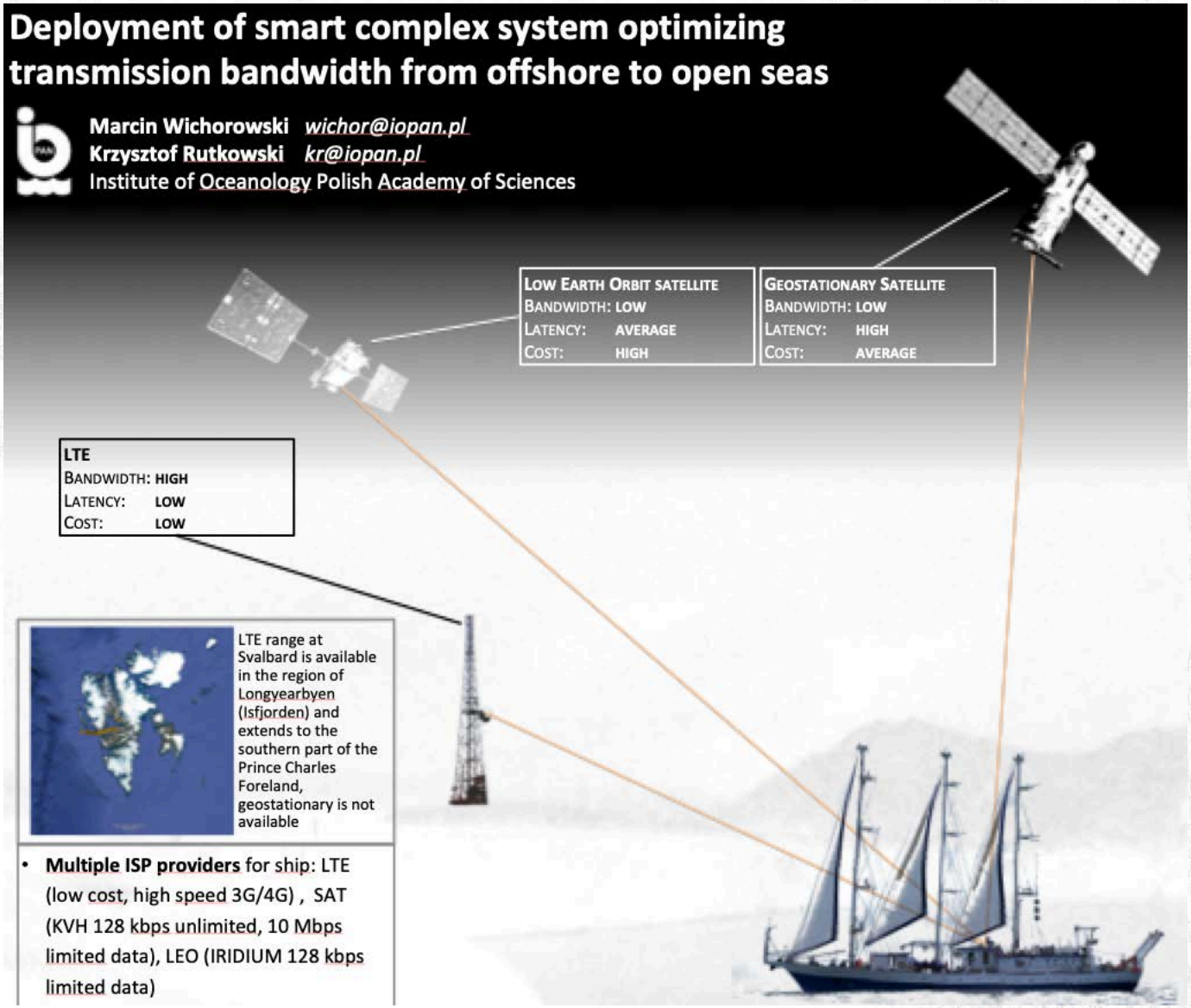
+ much more extended number of parameters

+ merged satellite and modeling products = everyday product

Variety and Velocity of the data streams

Deployment of smart complex system optimizing transmission bandwidth from offshore to open seas

Marcin Wichorowski *wichor@iopan.pl*
Krzysztof Rutkowski *kr@iopan.pl*
Institute of Oceanology Polish Academy of Sciences



LTE range at Svalbard is available in the region of Longyearbyen (Isfjorden) and extends to the southern part of the Prince Charles Foreland, geostationary is not available

- **Multiple ISP providers** for ship: LTE (low cost, high speed 3G/4G) , SAT (KVH 128 kbps unlimited, 10 Mbps limited data), LEO (IRIDIUM 128 kbps limited data)

eCUDO.pl – infrastructure for the assurance of the data FAIRness

„Elektroniczne Centrum Udostępniania Danych Oceanograficznych eCUDO.pl” is a Project funded within the frame of Operational Programme Digital Poland for 2014-2020, managed by Digital Poland Project Centre (CPPC) with allocated budget 3.5M EUR (84,63% –ERDF, 15,37% national budget).

The Project is carried out by consortium of 7 research and scientific institutions: Institute of Oceanology Polish Academy of Sciences (leader), Maritime Institute in Gdańsk, National Marine Fisheries Research Institute, Polish Geological Institute – National Research Institute, University of Gdańsk, Pomeranian University in Słupsk and University of Szczecin.

The goal: to develop Oceanographic Data and Information System providing unified access to national oceanographic science resources.

eCUDO.pl – infrastructure for the assurance of the data FAIRness

- 7 organisations
- 3.5M EUR budget
- up to 4 000 000 data objects to be digitized and open for community
- up to 500TB of data
- 5 * Open Data Model and FAIR assured data provision
- WCAG 2.0 conformed interface for human interaction
- M2M interface for interaction with outer systems

eCUDO.pl – infrastructure for the assurance of the data FAIRness

- **Findability**

data is provided as linked data and associated with metadata, propagated to outer systems (SeaDataCloud, EMODNet, BOOS)

- **Accesibility**

M2M interfaces use open protocols and data formats, well documented API's, metadata is provided by HA convergent infrastructure

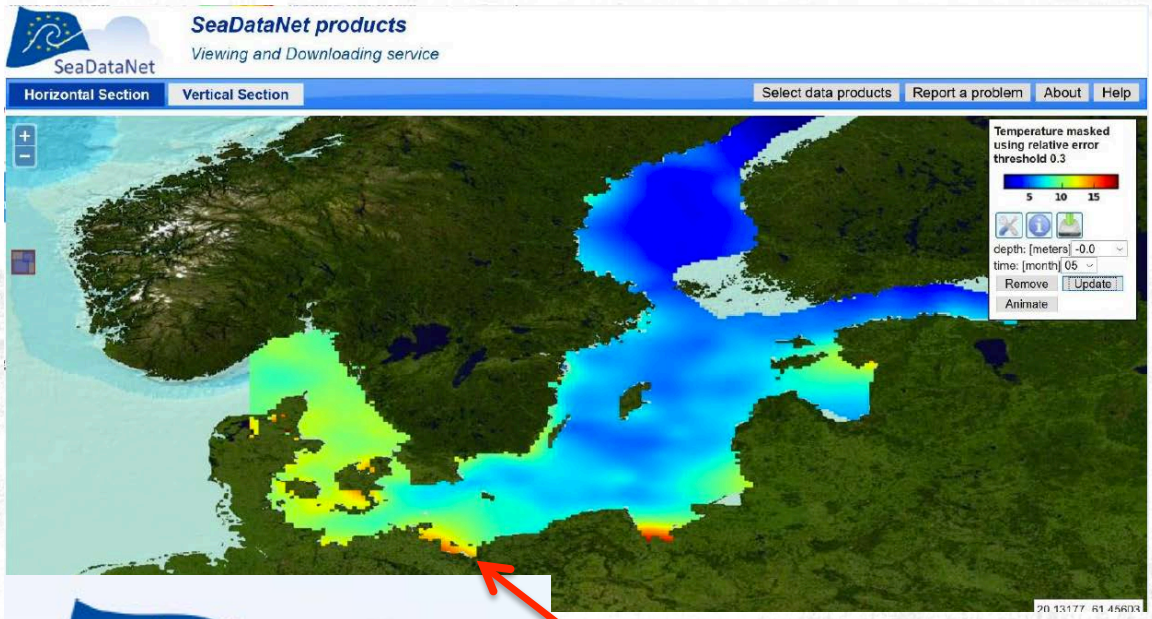
- **Interoperability**

identifiers of datasets are worldwide unique, metadata system base on BODC (SeaDataCloud) controlled vocabularies, data and metadata standards and formats are related to environmental sciences domain (INSPIRE, SeaDataCloud CDI, Darwin Core, NetCDF)

- **Reuse**

data is published on Creative Commons Attribution license

eCUDO.pl – interaction with community used systems



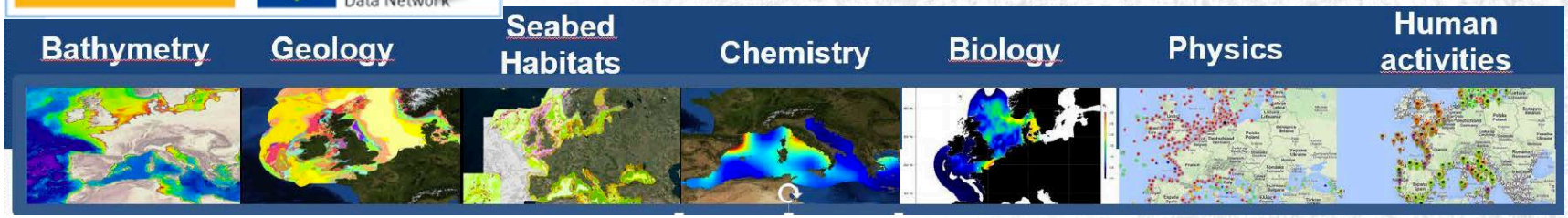
SeaDataCloud

eCUDO.pl
Elektroniczne Centrum Udostepniania Danych Oceanograficznych

ICES CIEM

EMODnet
European Marine Observation and Data Network

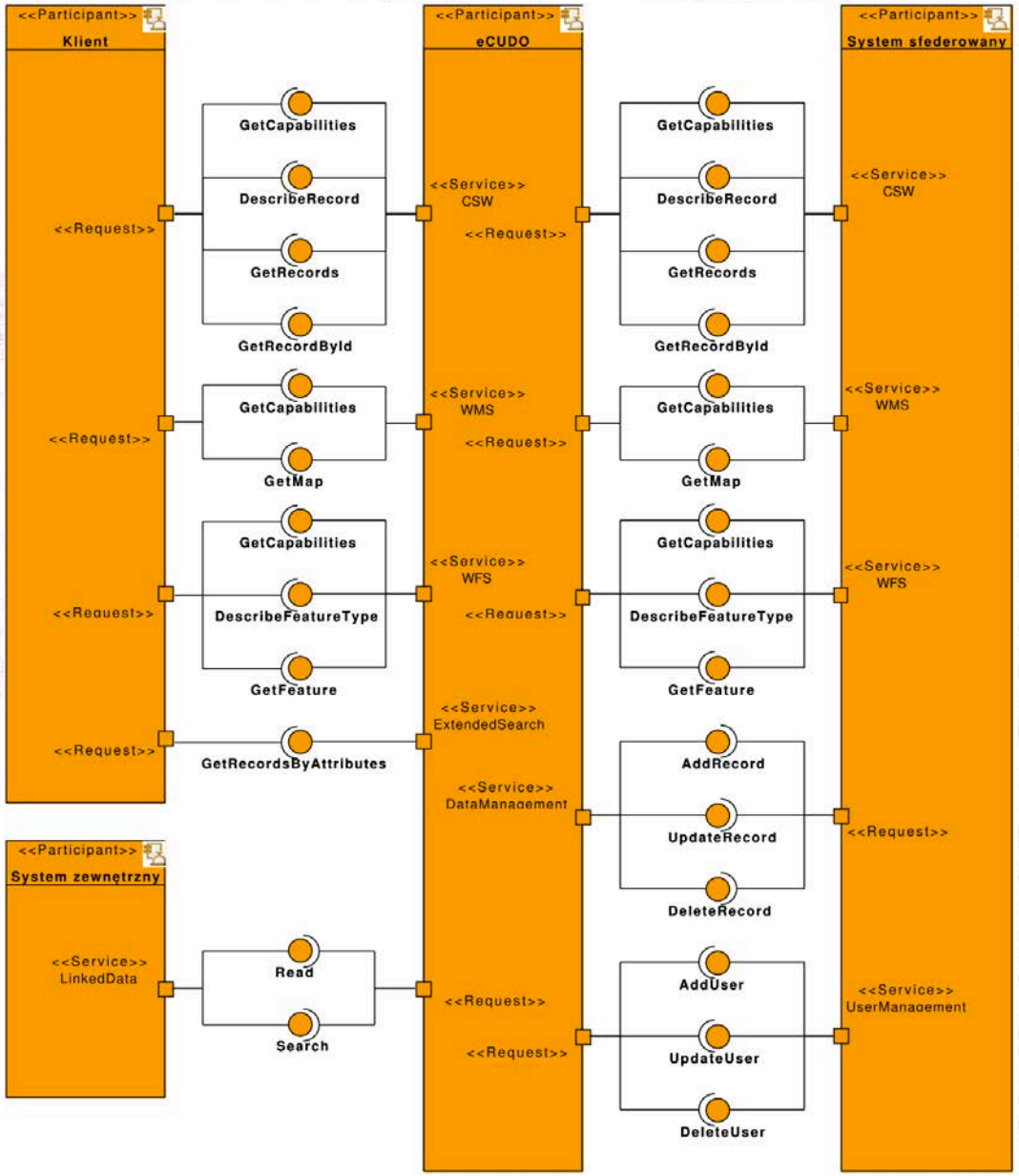
BOOS
Baltic Operational Oceanographic System



Assumptions of the design

- The central element of the platform is infrastructure based on a virtualization server connected to the data storage system. The use of such a solution allows for easy implementation of individual modules, as well as their replacement in the event of a need to conduct a complex software update or to remove errors in the operation of the system.
- The platform's infrastructure consists of a main server that provides required functionalities to external users, a single-sign-on server allowing one-time login to the platform and use of all its resources, spatial data server performing WMS, WFS and CSW services and a server processing data from non-federated external systems .
- The platform's environment (users, federated systems and non-federated external systems) can be distributed to any geographical location with communication via the Internet. For the sake of simplifying the model, it is assumed that each external system (both federated and un-federated) will consist of a server providing data, combined with a data warehouse.
- Users can access the system via any device (computer, laptop, tablet, mobile phone) with installed software supporting the display of websites using the http protocol or supporting WFS, WMS and CSW type services.

Distributed Oceanographic Data Centre



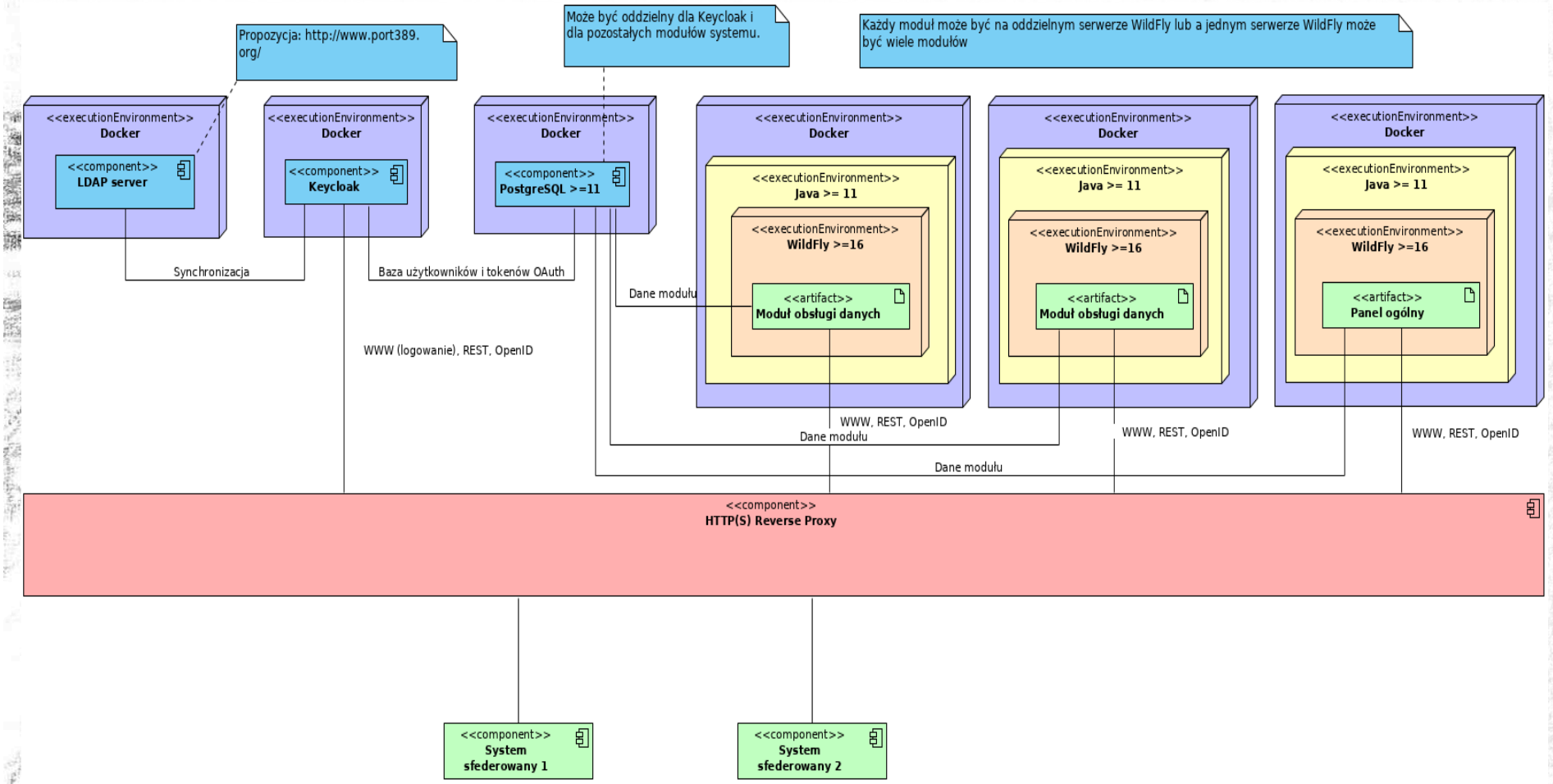
Assumptions of the design

- All virtualized system components work on the CentOS operating system.
- All applications running under the virtualization server operate as part of the virtual Java machine provided by the Java runtime environment for the CentOS system. In the case of web applications and enterprise-class applications, they will be launched as part of the WildFly application server running in the Java virtual machine.
- The PostgreSQL server is used as the database server with the PostGIS extension providing support for geographic data processing (GDAL).
- The central element is eCUDOApp, an enterprise class application consisting of the EJB eCUDOEngine module and the eCUDOWeb web module communicating with each other using local EJB interfaces. All operations on the database will be performed from the level of the EJB module through the JDBC protocol (Java Data Base Connectivity).
- A GeoServer application embedded in a Jetty container running in a Java virtual machine is used as a GIS data server. Both the GeoServer application and the Jetty container are free software available under the Open Source license. The application through the eCUDOWeb module will issue WMS, WFS and CSW services in the Representational State Transfer (REST) architecture. As in the case of eCUDOEngine, communication with the database will be implemented through the JDBC protocol.

Assumptions of the design

- The eCUDOLinkedDataParser application is launched directly in the Java virtual machine and in specified time intervals it will synchronize metadata from non-federated external systems using the http protocol (Hypertext Transfer Protocol) with metadata on the eCUDO platform via the eCUDOWeb module using the published services in REST and SOAP architecture (Simple Object Access Protocol). The application will be able to process data provided by external systems in the Linked Data format.
- The SSO server will contain the Keycloak web application acting as a central login point that allows the use of all platform elements after a single login. Keycloak is a free application available under the Open Source license. Login will be possible via the website by the http protocol as well as through authorization connected to REST and SOAP service calls.
- The above elements, the CentOS system, the WildFly application server and the PostgreSQL database server together with the PostGIS extension, are free software available under the Open Source license.
- The data presented by the eCUDOWeb module will come from external systems, especially from federated systems. Each of the federated systems will already have an enterprise application for storing and sharing data. A module or a separate application serving as a proxy for the central eCUDO system should be prepared for each of these systems.

Distributed Oceanographic Data Centre



Conclusions

Deployment of unified system providing access to data resources managed by key scientific organizations in Poland will trigger and added value for national economy through increase of data availability for all levels of administration, with concurrent decrease of the total cost of data acquisition, management and exchange.

Unified data formats and protocols will boost development of the services based on environmental data. Advanced services provided for clients (including data analysis services) extend availability of oceanographic data both to Polish and European organisations.

System is open for all stakeholders and ready to aggregate other organisations and data sources. Up to now cooperation has been agreed with Ministry of Environment Protection and Ministry of Maritime Economy and Inland Navigation – System for Spatial Information of Maritime Authorities, and Institute of Meteorology and Water Management

The eCUDO.pl project, when finished is intended to be well integrated with SeaDataCloud and EMODNet systems.

Further integration of operations on infrastructural and organisational level will provide solid basis for Polish Oceanographic Data Committee

Thank You!

Marcin Wichorowski, Institute of Oceanology Polish Academy of Sciences, ***wichor@iopan.pl***

Mirosława Ostrowska, Institute of Oceanology, Polish Academy of Sciences, ***ostra@iopan.pl***

Michał Piotrowski, Maritime Institute in Gdańsk, ***Michal.Piotrowski@im.gda.pl***

Lena Szymanek, National Marine Fisheries Research Institute, ***lszymanek@mir.gdynia.pl***

Urszula Pączek, Polish Geological Institute National Research Institute, ***upac@pig.gov.pl***

Michał Wójcik, Maritime Institute in Gdańsk, ***Michal.Wojcik@im.gda.pl***

谢谢

Fundusze Europejskie Polska Cyfrowa

Rzeczpospolita Polska

eCUDO.pl

Unia Europejska Europejski Fundusz Rozwoju Regionalnego