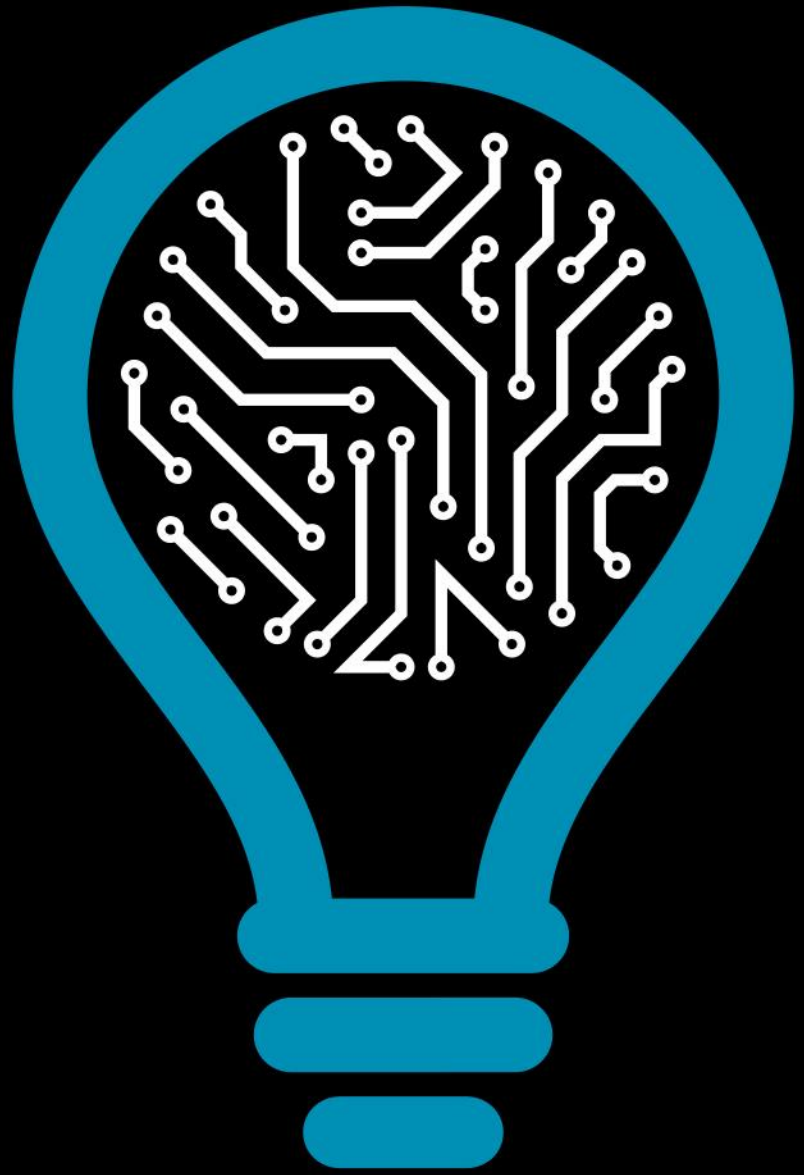
Da produção
de conhecimento
à inovação de
base científica



INESCTEC

INSTITUTO DE ENGENHARIA
DE SISTEMAS E COMPUTADORES,
TECNOLOGIA E CIÊNCIA

# Tools for Research Data Management: SeaBioData Case Study – INESC TEC & IPMA

SEA
BIO
DATA


Artur Rocha
Alexandre Costa
Carlos Almeida
Daniel Benevides
Inês Garganta
João Castro
José Alexandre Teixeira
Ricardo Amorim
Susana Barbosa
Cristina Ribeiro
Gabriel David


INESCTEC

ipma

# Project Summary

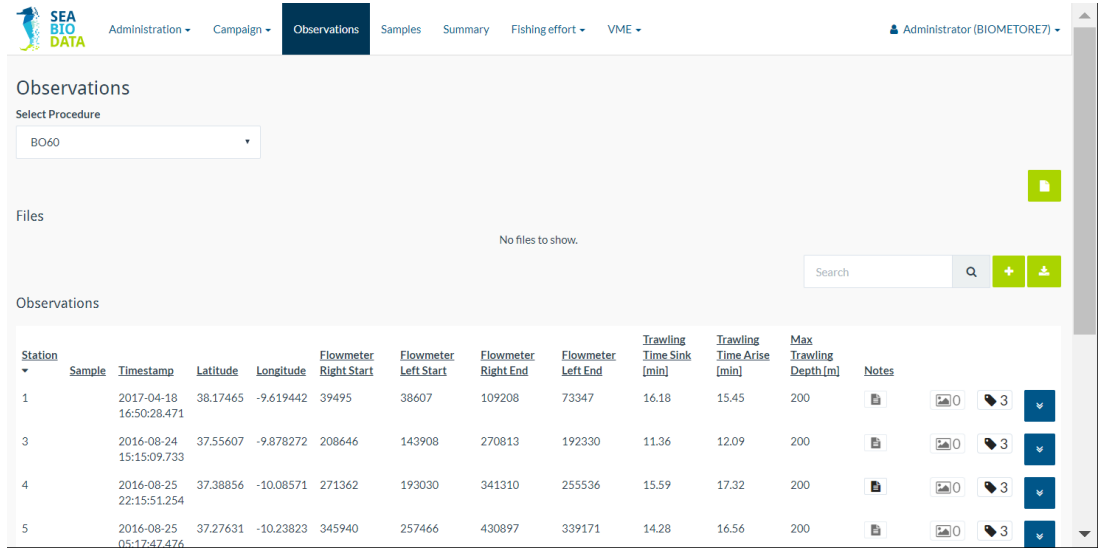
- **Title**: SeaBioData - Portuguese Seamounts Biodiversity Data Management
- **Reference**: PT02_Aviso5_0002
- **Partners**: INESC TEC and IPMA
  - consultant: IMR
- **Duration**: 21 months, 2015-07-16 / 2017-04-30
- **Funding agency**: EEA Grants
  - *OBJECTIVE EEA PT02 PA* Good Environmental State of European marine and coastal waters
  - Call 5 - National Ocean Data integration
- **Budget**: 229 K€

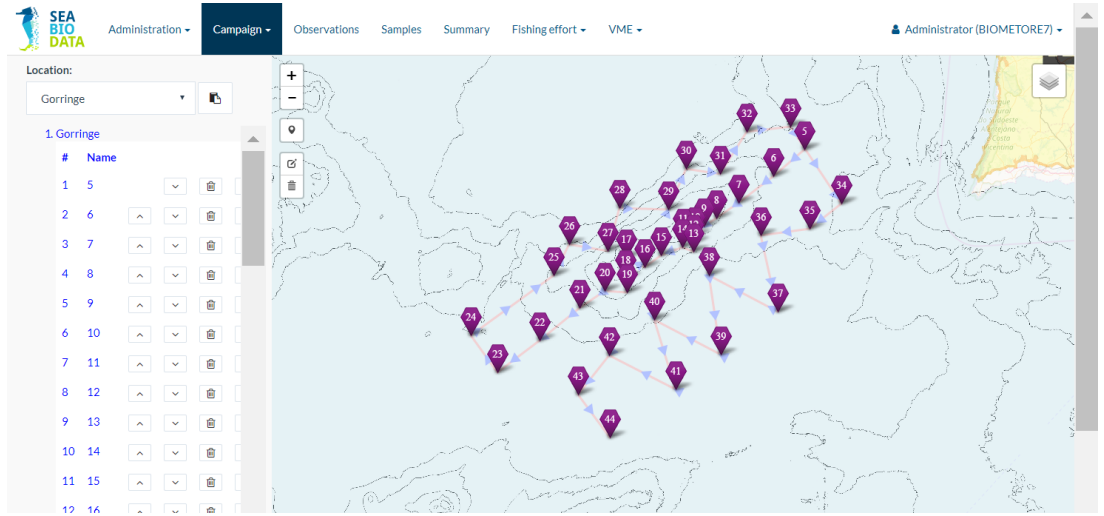
# SeaBioData Case Study:

# An Approach based on the DataBase Paradigm

# The Database Paradigm…

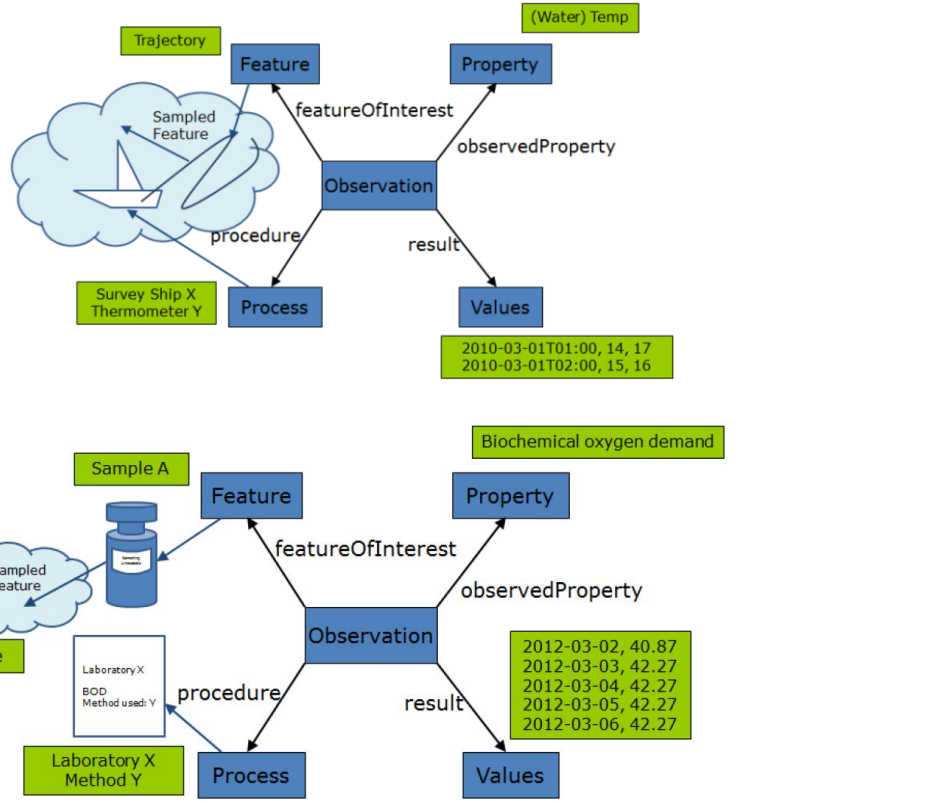
- Each “observation” corresponds to a line in the Database
- Enables:
  - SQL queries, relating different observations from different features
  - Sophisticated metadata
  - Elaborated value-added services
- Georreferenced data:
  - INSPIRE compliant (interoperable)
  - OGC Sensor Observation Service (ISSO 19153) compliant



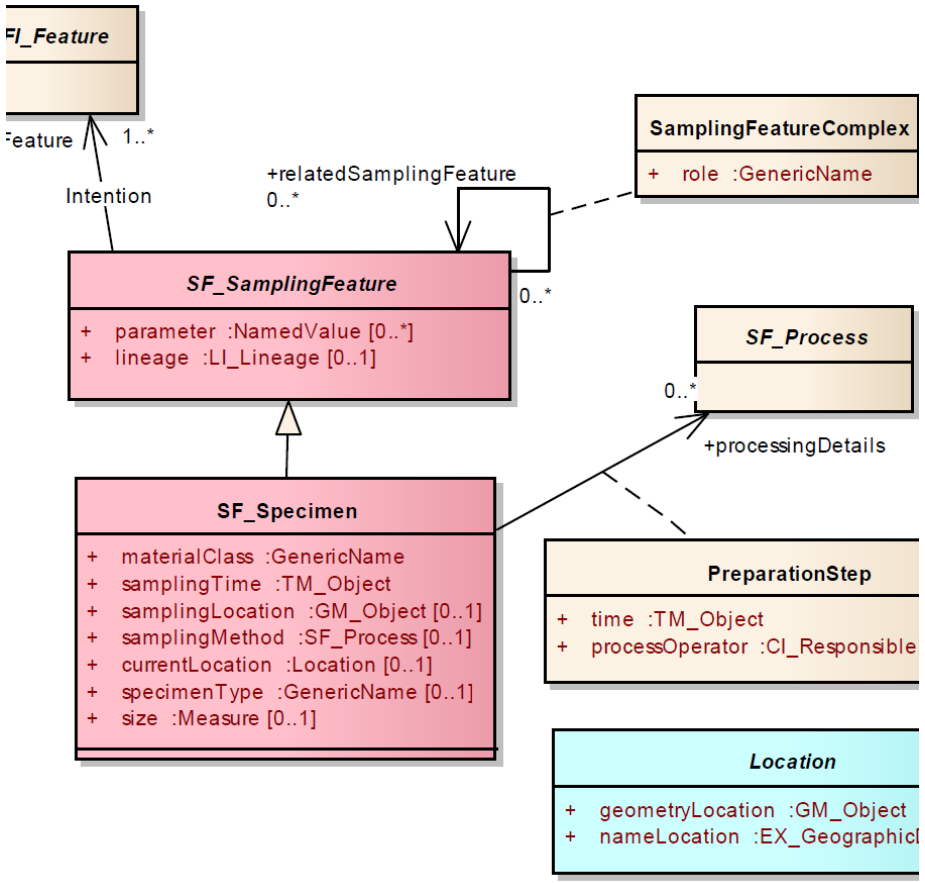
# Observation centric

- Feature of Interest (geographic feature)
  - Sampled Features
- Process (semi-structured)
  - Variety of procedures
- Observation (time, location, …)
- Observed Property
  - Simple types
  - Measurement with Units
  - Taxonomy, …
- Values



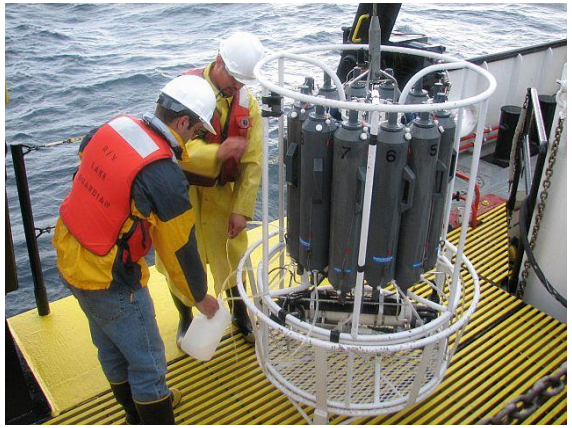
# Supports Complex Processes

- Possibility of relating several process steps in subsequent laboratory observations
- Concept of Sample
  - Water sample
  - Sediments
  - Specimen, …
- Sample Life Cycle
  - Preparation steps
  - Where is it deposited at
  - Collection ID



# Documenting Relevant Context Information

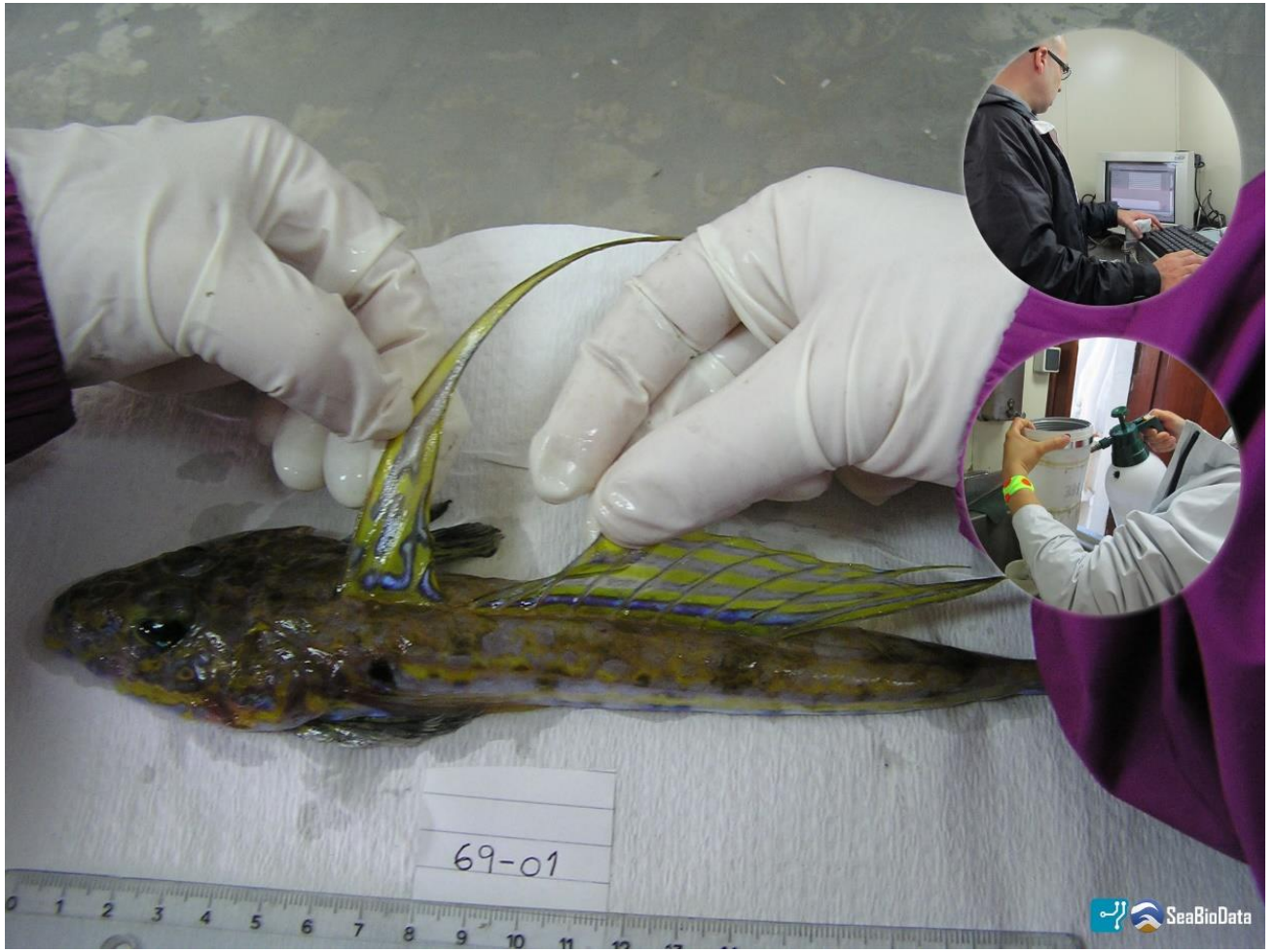
- Projects
- Surveys
- Equipment
- Teams
- Stations/Locations
- …



biometore

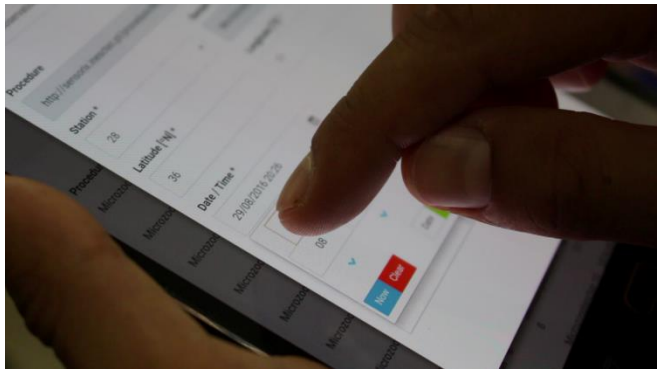
# Repository of Data and Annex Documents

- Photos, vídeos, sounds, …
- Survery reports
- Instrument outputs (e.g. CTD)
- Relation to external repositories (e.g. samples deposited at MUHNAC)
- Derived outputs
- …
- Files are associated to the respective concepts in the data model



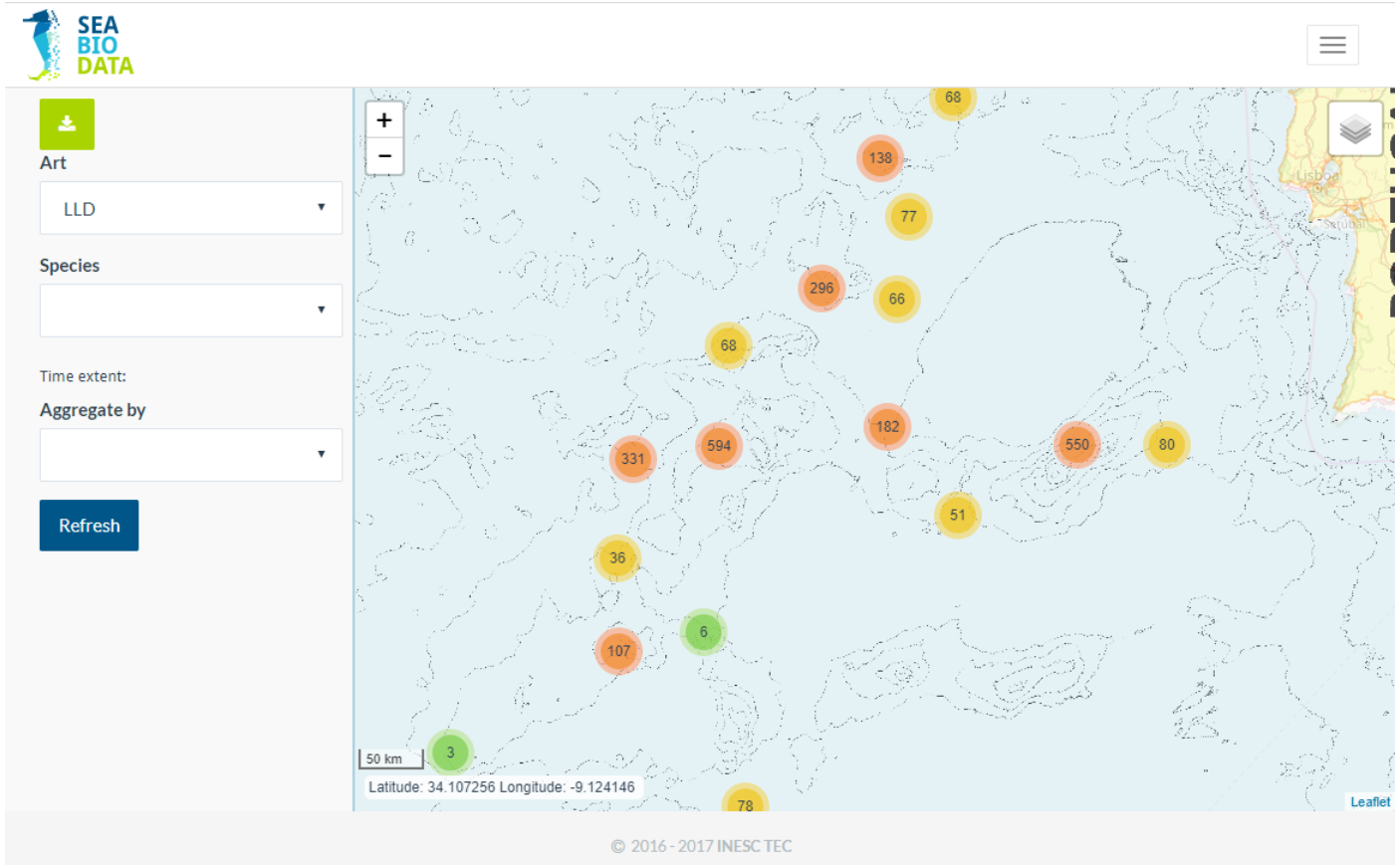
# Promoting Innovative Ways of Collecting Data

- Using Tablets as Log Book
- Flexible forms generated according to the Observation processes
- Filling aids



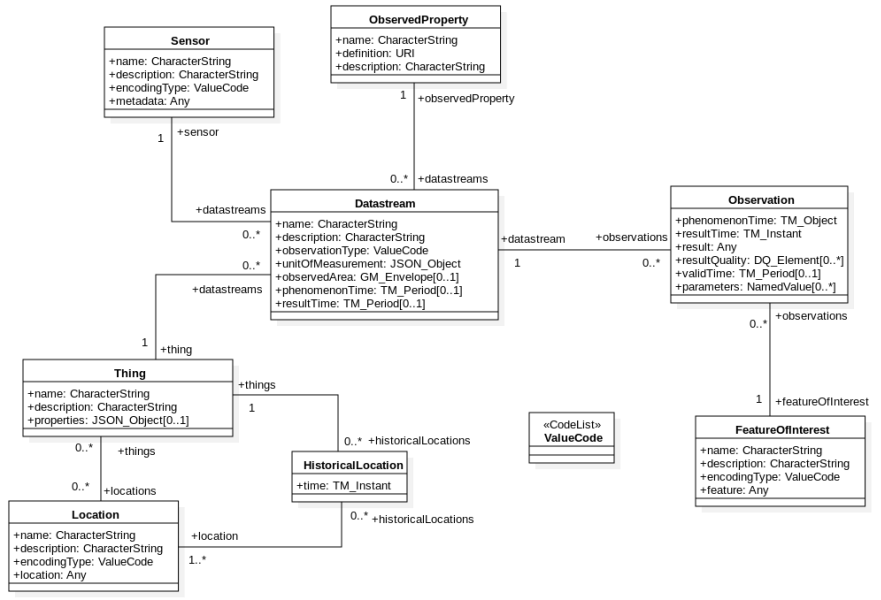
# Value added services

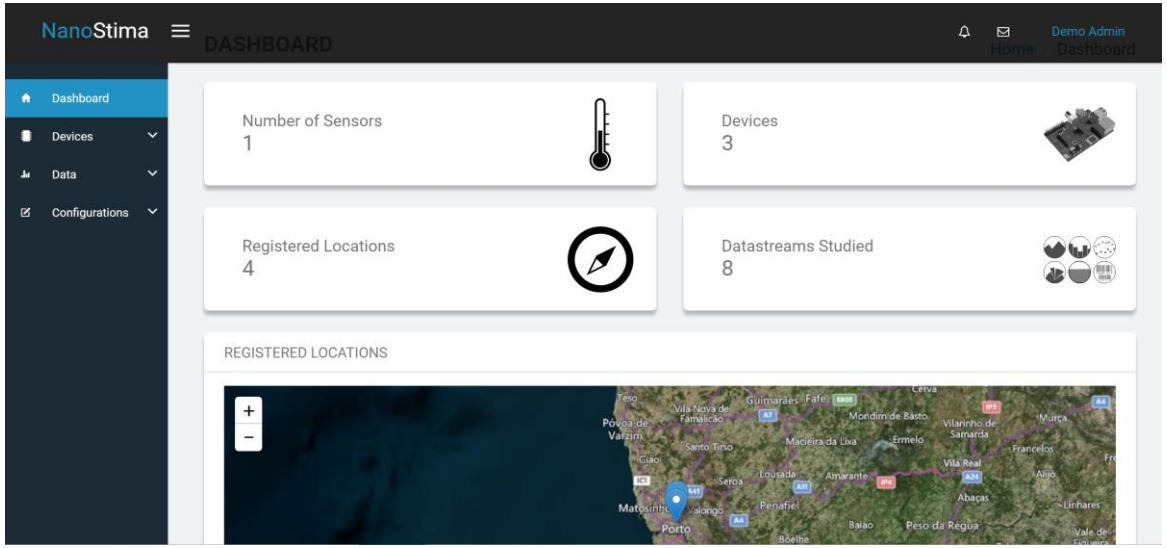
- Human pressures (fishing effort, in picture)
  - Vessel Monitoring System
  - Fishing Logs (“diários de pesca”)
- Vulnerable Marine Ecosystems



# Current trends – IoT: Real Time Data

- Common data model for “data streams”
- The OGC SensorThings API (sensing profile) is derived from the SOS conceptual model
- Restful API
- JSON encoded



# Consumable by Generic Clients

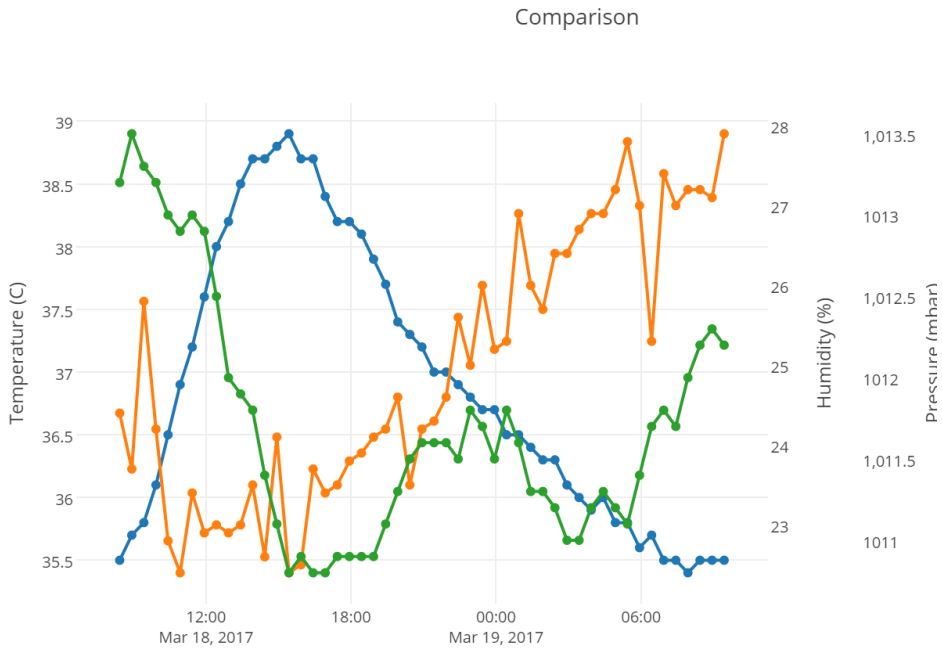
Example in picture: data streams at INESC TEC consumed by online client from SensorUp (Canada)

SensorThings Share

sensorup

Templates ▾ Tools ▾ Share ▾ Run

```
<!DOCTYPE html>
<html>
<head>
  <meta charset='utf-8' />
  <meta name='viewport' content='initial-scale=1,maximum-scale=1,user-scalable=no' />
  <link href='https://fonts.googleapis.com/css?family=Open+Sans:400,300,700' rel='styles
  <script src="https://cdn.plot.ly/plotly-1.22.0.min.js"></script>
  <script src="http://sdk.sensorup.com/sta-chart/sta-chart-0.0.8.min.js"></script>
  <style>
    body {
        font-family: 'Open Sans', sans-serif, Arial;
        font-size: 12px;
    }
  </style>
</head>
<body>
  <div id="vis"></div>
  <script type="text/javascript">
  stachart.generateChart({
    'staBaseUrl': 'http://nanostima-rl3-2-2.inescn.pt:8080/sensorthings/noauth/v1.0',
    'datastreamIds': [ '8', '9', '10' ],
    'observations': {
      retrieveAll: true,
      timeRangeCode: 'last',
      timeRangeAmount: 3,
      timeRangeUnit: 'd'
    },
    'titleTemplate': 'Comparison',
    'plotlyData': [
      { name: 'Temperature', yaxis: 'y' },
      { name: 'Humidity', yaxis: 'y2' },
      { name: 'Pressure', yaxis: 'y3' }
    ],
```



# Conclusions & Future Work

- The Database paradigm for storing research data is complementary to other paradigms:
  - exporting a **selection of observations**
  - to **file-based datasets**
  - for
    - Dataset exchange
    - Publishing
- Issues:
  - Citing *open* (evolving) data sets
  - More value-added services (e.g. visualization, exploration)

# SeaBioData
# Live Demo