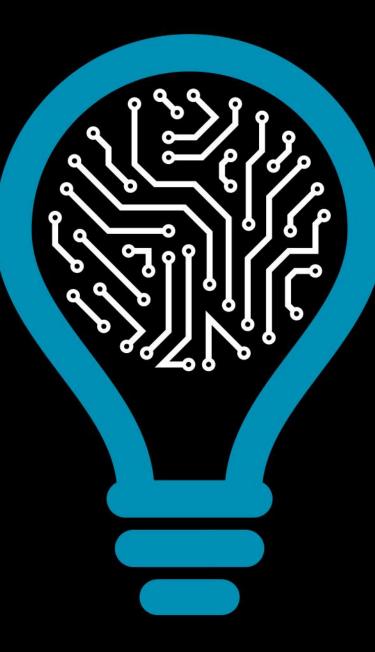
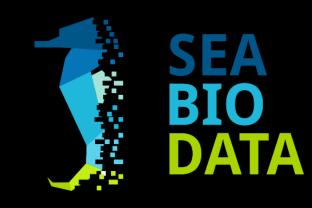
Da produção de conhecimento à inovação de base científica





INSTITUTO DE ENGENHARIA DE SISTEMAS E COMPUTADORES, TECNOLOGIA E CIÊNCIA

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







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

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 PAGood Environmental State of European marine and coastal waters
 - Call 5 National Ocean Data integration
- **Budget**: 229 K€





3



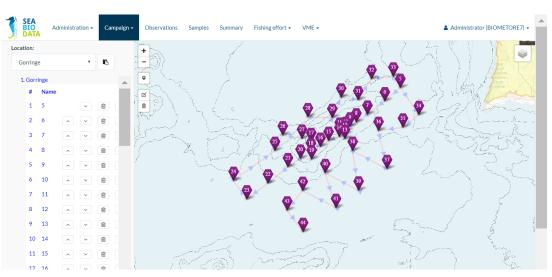
SeaBioData Case Study:

An Approach based on the DataBase Paradigm

The Database Paradigm...

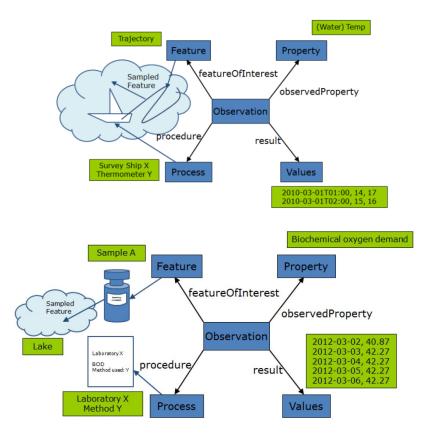
- Each "observation" corresponds to a line in the Database
- Enables:
 - SQL queries, relating different observations from different features
 - Sophisticated metadata
 - Elaborated value-added services
- Georreferenced data:
 - INSPIRE compliant (interoperable)
 - OGC Sensor Observation Service (ISSO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

SEA BIO DAT		Administration +	Campai	ign - Obs	ervations	Samples Sum	mary Fishing	effort 🗸 🛛 VMI	E +			🛔 Ad	ministrator	(BIOMET	ORE7) 🗸
Observ elect Proce		ns													
BO60			•												_
iles							No files to	chow							
							NO TITES LO	SHOW.			Search		c	λ +	±
bservati	ons														
Station Sa	ample	Timestamp	Latitude	Longitude	Flowmeter Right Start		Flowmeter Right End	<u>Flowmeter</u> Left End	<u>Trawling</u> <u>Time Sink</u> [<u>min]</u>	<u>Trawling</u> <u>Time Arise</u> [<u>min]</u>	<u>Max</u> <u>Trawling</u> <u>Depth [m]</u>	Notes			
		2017-04-18 16:50:28.471	38.17465	-9.619442	39495	38607	109208	73347	16.18	15.45	200	È	0	₿3	*
3		2016-08-24 15:15:09.733	37.55607	-9.878272	208646	143908	270813	192330	11.36	12.09	200	È	•	\$ 3	*
Ļ		2016-08-25 22:15:51.254	37.38856	-10.08571	271362	193030	341310	255536	15.59	17.32	200	8	•	\$ 3	*
5		2016-08-25	37.27631	-10.23823	345940	257466	430897	339171	14.28	16.56	200	Đ	1 0	\$ 3	¥



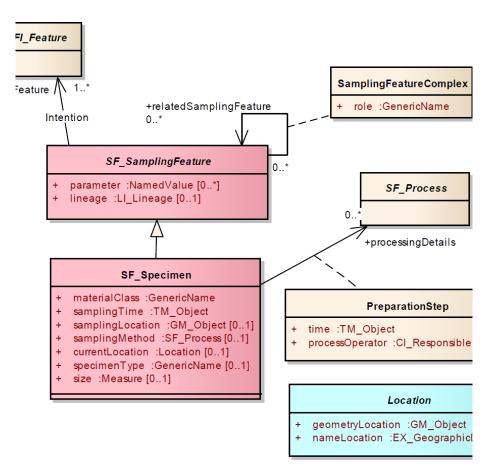
Observation centric

- 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



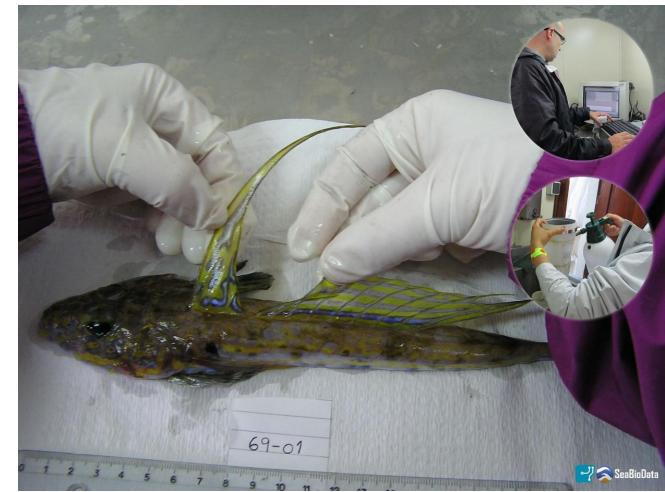




biometore

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

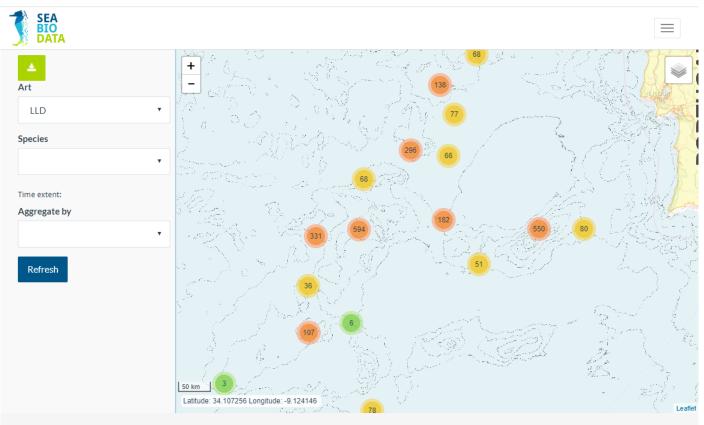
- Using Tablets as Log Book
- Flexible forms generated according to the Observation processes
- Filling aids





Value added services

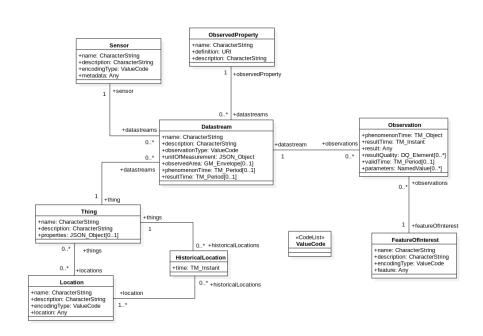
- Human pressures (fishing effort, in picture)
 - Vessel Monitoring System
 - Fishing Logs ("diários de pesca")
- Vulnerable Marine Ecosystems

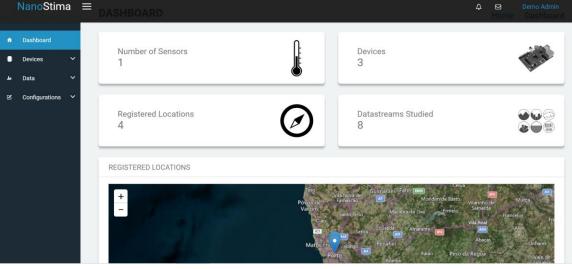


© 2016 - 2017 INESC TEC

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 Comparison 1 <!DOCTYPE html> 2 - <html> 3 - <head> 4 <meta charset='utf-8' /> <meta name='viewport' content='initial-scale=1,maximum-scale=1,user-scalable=no' /> 5 39 <link href='https://fonts.googleapis.com/css?family=Open+Sans:400,300,700' rel='styles</pre> 28 6 1.013.5 7 <script src="https://cdn.plot.ly/plotly-1.22.0.min.js"></script> 8 <script src="http://sdk.sensorup.com/sta-chart/sta-chart-0.0.8.min.js"></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script></script 38.5 9 -<style> 27 1013 10 body { font-family: 'Open Sans', sans-serif, Arial; 11 38 12 font-size: 12px; 13 } 26 Temperature (C) 1,012.5 14 </style> umidity (%) 37.5 15 </head> 16 - <body> 25 <div id="vis"></div> 17 37 1012 <script type="text/javascript"> 18 -19 stachart.generateChart({ 20 'staBaseUrl': 'http://nanostima-rl3-2-2.inescn.pt:8080/sensorthings/noauth/v1.0', 36.5 24 21 'datastreamIds': ['8', '9', '10'], 1,011.5 22 -'observations': { 36 23 retrieveAll: true, 24 timeRangeCode: 'last', 23 1011 25 timeRangeAmount: 3, 35.5 26 timeRangeUnit: 'd' 27 }, 28 'titleTemplate': 'Comparison', 12:00 18:00 00:00 06:00 29 -'plotlyData': [Mar 18, 2017 Mar 19, 2017 30 { name: 'Temperature', yaxis: 'y' }, 31 { name: 'Humidity', yaxis: 'y2' }, 32 { name: 'Pressure', yaxis: 'y3' } 33]; 24

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