

Towards the LifeWatch Big Data Model construction (II)

10th LW Stakeholders Board, Crete, 7-8th July Session: Moving to fast initial construction



presented by Jesus Marco (<u>marco@ifca.unican.es</u>) Instituto de Física de Cantabria (IFCA) CSIC, NATIONAL RESEARCH COUNCIL SANTANDER, SPAIN



ICT-Core Starting Tasks

ICT CORE	Start-up activities		
Keep Reference Model up to date	Mechanism will be developed. Currently		
	expansion done by ENVRI and EUDAT.		
Analysis of requirements	Need requirements from distributed initiatives.		
ICT-core technical unit project plan	Proposal will follow with lean organization with		
	coordination and outsourcing capabilities.		
Technical framework user portal	Priority for e-science users' portal. Cloud/Grid		
	experiences will assist in drafting proposals.		
IT release plan and annual work plans	Will follow (after tasks 26 and 27)		
Core basic Application Services	Priorities of core basic application services for		
	the initial years to be proposed.		
Organize distributed	A management tool will come into place to keep		
construction/operations	track of distributed activities and relations		
	related to the distributed e-Infrastructure		
	construction/operations.		
	A technical body will be created.		
Contribute to arrangements with data	Test cases to be addressed ((in cooperation with		
resources	EUDAT, ENVRI, EUBON, LTER and GBIF).		
Contribute to enabling data generation	Sensor enabled data generation is being		
	addressed.		



LifeWatch 10th Stakeholders Board

Solution for HETEROGENEITY: An SOA approach



How to explore the LW Core-ICT Implementation As presented at Interministerial in Seville (July 2013)

A SUGGESTED PATH:

Revise Key Components and Actors

- Learn from Preparatory Phase and from on-going projects
- Learn from other Research Infrastructures
- Interact with all partners in LW
 - Learn, collaborate, build relationships
- **IN ORDER TO CONTRIBUTE TO A REVISED TASK LIST**



A pilot project to understand the global framewo *Adaptation and improvement of the e-Infrastructure ICTS-EBD (Estacion Biologica de Doñana)*

- Funded by MINECO (CSIC to be commissioned to execute it starting in 2013)
 - Setup an operational framework supporting from basic services to advanced data processing and collaborative work
 - Improve the sensor monitoring network at Doñana
- MATCH & INTEGRATE ICT Services CAPABILITIES IN ANDALUCIA

Progress with LW Core-ICT Implementation

- A pilot project to understand the global framework: Adaptation and improvement of the e-Infrastructure ICTS-EBD
 - Funded by MINECO (CSIC to be commissioned to execute it)
 - Setup an operational framework supporting from basic services to advanced data processing and collaborative work
 - Improve the sensor monitoring network at Doñana
 - MATCH & INTEGRATE ICT Services
- Four actions:
 - ICT e-Infrastructure
 - servers (cloud oriented + HTC/HPC), storage (300TB), 10GB network
 - methodology for services design
 - Implementing services
 - Integration of the network of sensors
 - Conditioning for in-situ data process at RBD (Natural Reserve)
 - New sensors and observational setups

in progress!

before end of 2014



Reflection on our context

- LW "global" funding is limited
 - **Focus on coordination + selected global services**
- 1- National initiatives/results must be integrated
- 2- Coordinate with EU/Global initiatives with resources:
 - E-Infrastructures: EGI, EUDAT, PRACE
 - Data: GBIF, LTER

CSIC

- OTHER ESFRI Initiatives
- 3- Exploit previous/ongoing results from EU projects
- 4- Consider new H2020 opportunities
- 5- Can we engage SMEs/Industry?
- 6- What about Public Managers?
- 7- Can we support Citizen Science?





EGI services for LW:

in progress!

- MODEL: LW brings users & resources together!
 - LW core-ICT (Spain) will operate an e-infrastructure in 2014
 - in progress! LW core-ICT could/will integrate grid/cloud infrastructure
 - LW VOMS will be supported by LW core-ICT
 - LW core-ICT will rely on IberGrid for this integration in EGI
 - LW national initiatives will be integrated
 - LW core-ICT will support integration at different levels (NGI role?)
 - **LW** will explore successful examples in EGI FedCloud:
 - EUBrazilOpenBio Ecological Niche Modeling Service
 - FUBrazilCloudConnect

CSIC

- New challenge for phenology with LTER/Univ.Granada
- So, LW will use existing EGI services

EUDAT services for LW?

- EUDAT has two key sides for LW:
 - Knowledge about DATA management - ¦H
 - New services: - !н
 - B2 SHARF
 - B2 STAGE
 - B2 FIND
- in progress! LW can/should explore them! Contribute to Proposals Call (befor 26th Feb!)
- Select topic(s):
 - Real Time -5-
 - Semantic Mapping -, н
 - Workflow Execution -___
 - Data Lifecycle ╌╬┥



Use case: Research Vessel Simon Stevin



Additional services

- Additional services are being studied:
 - Considering also output of ongoing projects
 - ENVRI, BIOVEL, COOPEUS, iMARINE, CREATIVE-B



- Some of them:
 - Identity federation for researchers, educators and students
 - Digital Identifier e-Infrastructure for digital objects (and PID issues)
 - Simple Storage/File System + Medium/Large DBMS cloud/grid instances
 - Large, persistent DBMS, GIS systems in cloud/grid framework
 - Parallel (multithread?) datamining (in phytom OPR) cloud/grid instance
 - Systems to handle & process real time heams
 - Access to large databases/directives common to other research areas
 - Workflows connecting to HPC resources (o(10²-10³) processes, 1-100 TB)
 - Support to virtual eLaboratory
 - Data discovery and access

Along 2014 we need to work to complete a VRE proposal



H2020 opportunities

- EINFRA-2014-1: e-Infrastructure for Open Access
 - Community Knowledge Tool
- EINFRA-2014-2:
 - Manage/Preserve Big Research Data
 - RDA
 - HPC
- INFRAIA-2014-2015:
 - support new communities (ex. LTER sites? NETLAKES?...)
- INFRAIA-2015-1: New skills
- INFRASUPP-2014-2: International Collaboration
- INFRADEV-1-2014: ESFRI Clusters
- INFRADEV-3-2015: ESFRI projects operation
- ICT-2014-1: Cloud
- ICT-2014-2: 5G Network



LifeWatch 10th Stakeholders Board





Can we address a challenge?

- Grand Challenge: Predictive Modeling of Biosphere in progress!
 - Global Carbon cycle JOIN us at EGU meeting (29 April,)
 - Essential Biodiversity Variables (EBV) for IPBES
 - IPBES=Intergovernmental Platform on Biodiversity & Ecosystem Services (cf. IPCC)
 - EBV= a measurement required for study, reporting, and management of biodiversity change.
 - Examples of candidate EBV:
 - Species populations: Abundances and distributions (inc. invasive alien)

		EXAMPLES OF CAN	DIDALE E	SSENTIAL DIODIVERSITY V	AKIADLES	
EBV class	EBV examples	Measurement and scalability	Temporal sensitivity	Feasibility	Relevance for CBD targets and indicators (1,9)	
Genetic composition	Allelic diversity	Genotypes of selected species (e.g., endangered, domesticated) at representative locations.	Generation time	Data available for many species and for several locations, but little global systematic sampling.	Targets: 12, 13. Indicators: Trends in genetic diversity of selected species and of domesticated animals and cultivated plants; RU.	Pereira et al., Science 2013
Species populations	Abundances and distributions	Counts or presence surveys for groups of species easy to monitor or important for ES, over an extensive network of sites, complemented with incidental data.	1 to >10 years	Standardized counts under way for some taxa but geographically restricted. Presence data collected for more taxa. Ongoing data integration efforts (Global Biodiversity Information Facility, Map of Life).	Targets: 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15. Indicators: LPI; WBI; RLI; population and extinction risk trends of target species, forest specialists in forests under restoration, and species that provide ES; trends in invasive alien species; trends in climatic impacts on populations.	
Species traits	Phenology	Timing of leaf coloration by RS, with in situ validation.	1 year	Several ongoing initiatives (Phenological Eyes Network, PhenoCam, etc.)	Targets: 10, 15. Indicators: Trends in extent and rate of shifts of boundaries of vulnerable ecosystems.	
Community	Taxonomic	Consistent multitaxa surveys and	5 to >10	Ongoing at intensive monitoring sites	Targets: 8, 10, 14.	

How to move?

- Next presentations will show more information needed/discussion
- Tomorrow discussion
- Plan for basic start setup
 - Minimal central services
 - Integrating existing national services
 - Examples: LW Sweden, LW Belgium
 - ...
 - Coordination with
 - Italy Service Center, Netherlands
 - EGI, EUDAT, PRACE
 - GBIF, LTER
 - Other ESFRI
 - International Collaborations
- Organization:
 - Task teams? IC3?...
- Knowledge map
 - FINAL USERS
 - EXISTING WORK AND EXPERTISE AT ICT-BIODIVERSITY LEVEL



LifeWatch 10th Stakeholders Board



JOINT RESEARCH UNIT LIFEWATCH SPAIN (JRU LW.ES) GOVERNANCE SCHEME

