LifeWatch e-Science European Infrastructure for Biodiversity and Ecosystem Research



### ATTIVITA' IN CORSO & PRIORITA' STRATEGICHE Gruppo di lavoro ICT

Paolo Tagliolato, Alessandro Oggioni, Nicola Fiore

<u>con Paolo Colangelo (CT Mediterraneo)</u>



# Virtual Laboratories

- Virtual Lab: remotely accessible virtual environment offering interactive opportunities to collaborate with related web services.
- Web service: access to web-based applications through a standardized interface allowing different applications from different sources to communicate each other.
- Workflow: series of activities necessary to complete a task -> an orchestrated sequence of data transmissions.

(Wouter Los, LifeWatch Vlab meeting, Amsterdam, March 2015)





### "Workflow"







# Formalization -> implementation

Use of Scientific workflow systems -> formalization leads to runnable implementation



Advantages:

- Design comparability
- Results comparability
- Local Reproducibility





## Service Oriented Science

Embed in an e-infrastructure -> performant web-services **orchestration** 





The workflow as a runnable formalization of scientific work

#### Taverna workflow system

- Open world assumption
- Composed by "services" (remote and local) configurable assignment to agents
- Workflow execution: data driven (plus control links)
- Nested workflows (nested Petri Nets)





#### Example workflow: alien species

# Workflows as VLab tools?

- Questioning workflows as tools for (not only IT-) scientists
  - Which users for VLabs?
    - IT skilled (with different skills: e.g. Engineers, Computer Scientists, Statisticians, bioinformatics experts, biologists with ecoinformatics skills ...)
    - Less IT skilled (e.g. biologists, ecologists, tassonomists...) and not programmers.
  - Which **roles** have users (both practically and theoretically) with respect to workflows?





## The user as an architect

- At **design time** the user can operate on an (even existing) workflow making some tuning.
- He can for example change some WF service parameter.
- Even simple operations require some knowledge of the workflow system.





## The user as an *agent*

- *Real life* Scientific workflows include processes acted both by computers and by researchers.
- Modeling the **user as an** *agent* **of the workflow execution** is a key point within the formalization of workflows.
- Example: any human choice occurring at runtime in dependence of some other task result

Further advantages:

- Improving workflow reusability (not only reproducibility)
- The workflow as a VLab *tool* (not only as a runnable formalization of past analyses).





## Implementation Aims

- Enhance efficiency
- Offer users tools for daily tasks (not only Taverna services: web services; web GUI apps)
- Offer (non IT) scientists interfaces they can actually use (developing for the user as an agent)
- Develop for architects reusable components...
- ...to create friendly workflows for the daily work of others users!





#### User interaction

• User as an agent in Taverna:

#### $\rightarrow$ Simple predefined interaction services



🏙 Nested workflow - A service that allows vou to have one workflow nested within and

#### Let's Hack it!





#### Enhancing the agent-user interaction

Progresses: reusable web-application interaction services with advanced interactivity

Eunis Species Names from Thesaurus (via SPARQL endpoint)	
Type some character of a species name and select one from the suggestions	Species list
primulae	Acyrthosiphon primulae Oligonychus primulae
Acyrthosiphon primulae	
Chromatomyia primulae	
Eusphalerum <b>primulae</b>	
Oligonychus primulae	

Some example:

- (linked data) Compose list of species names from Eunis Thesaurus (controlled vocabulary) via SPARQL endpoint
- Parameterization of several WF services within a single user interaction
- Exposition of workflow results in a web report

Example: (linked data) Eunis Species list composer

The interaction service was developed in the BioVeL project





### Enhancing the agent-user interaction

- Progresses: reusable web-application interaction services with advanced interactivity and processing
- Taverna aware Shiny (R) applications



#### Biotic/Abiotic/Pressure variables extraction



The interaction service was developed in the BioVeL project

Some example:

- Enrich species occurrence with abiotic data from WorldClim
- Choices of parameter through *descriptive statistics*
- Dynamically reshape data

Example: species occurrence enrichment with WorldClim abiotic data







#### **Re-combination**



#### Data Access - Vlab/data portal integration



https://dx.doi.org/10.6084/m9.figshare.1597713





#### LifeWatch

e-Science European Infrastructure for Biodiversity and Ecosystem Research



#### Thanks

