An Environment Supporting the Production of Live Research Objects
“Publishing data in a reusable form to support findings must be mandatory” [Science as an Open Enterprise, The Royal Society]

- all the elements exploited (primarily grey elements) are not available or not linked to the scientific result;
- It makes difficult to completely understand the results;
- It makes difficult to validate the results.
Live Research objects in a nutshell

- An abstraction for communicating, sharing and reusing research results:
  - aggregate all the “pieces” that contribute to a research result.

How can all these pieces stand together?

How can you produce live research objects?
The idea: producing live research object/ Virtual workspace

Remote workspace
- core element
- resemble classical folder based file system
- containing every item needed for creating live research objects

Sharing

Datasets
Workflows
Tech. Reports
Images, Videos

GL14 - M. Assante et al. An Environment Supporting the Production of Live Research Objects - 29/30 November, Rome
Producing live research object/ **Editing phase**

- **static structured text**
  - sections
  - headings
  - paragraphs

- **static or dynamic image** e.g.
  - result of a computation
  - satellite periodic measurements

- **tabular data**

- **tabular data representation**

"filtered" quantitative data set (with rich metadata and provenance information) e.g.
  - observations
  - statistics
  - time series
Producing live research object/ collaborative work

People to work collaboratively to the creation
- drafting
- writing
- assembling
- reviewing

Workflow governing the production
- several iterative phases (steps)
- responsible person and permissions
- manual and/or automatic routing
- concurrent access
Production of live research objects

- **Virtual Workspace**
  - from binary files to compound information objects

- **Editing framework**
  - define the structure of a live research object
  - entering content and compile them

- **Workflow Engine**
  - define the workflow governing their production
  - specifying the phases and users
Users can organise and share very different items

- tabular data, species distribution maps, time series

Sharing

Smart Folders

Desktop Access
Defining the structure

- define the structure of a live research object
- component oriented approach (static & dynamic)
Compiling a Live research object

- compliant with one of the defined templates
- complete or instantiate the dynamic components
Workflow Engine – defining a workflow

- Work collaboratively to the creation of a live research object
- Define the workflow governing the production of a live research object
- Reuse
Specify the steps and the relative responsible actor(s)

3 phases

**WF Template**
- author

**R. Object draft**
- .......
- ........
- ...........

**STEP 1**
- ROLE 1
- ROLE 3
  - Read
  - Read Update
  - Delete

**STEP N**
- ROLE 1
- ROLE 5
  - Read Update
  - Read Update Comment

**R. Object draft to Workflow Template Association**

**Roles to Permissions Association**

**Roles to Users Association**
exploits the capabilities offered by an underlying *Data Infrastructure* (DI).

[*digital infrastructure for data sharing and consumption]*

The *gCube software system*, whose technological development has been coordinated by ISTI-CNR and funded by E.C implements the DI approach.

- operates a **large federation** of computational and storage resources;
- equipped with **software frameworks** for data management;
- **supported data types** cover a wide spectrum ranging from **tabular data** to **research products**.
Production of Live Research Objects aimed at estimating the probability of marine species distribution in a global scale:

- some descriptive text
- data on the species gathered from authoritative data sources
- environmental data reporting on ecological elements
- algorithms aimed at estimating the probability of the occurrence of a species in a given area
- images of maps resulting from the algorithm(s)
Conclusion

- A comprehensive framework supporting the entire lifecycle of Live Research Objects production and management

- It has been designed and implemented in the context of two successive EU projects:
  - D4Science-II (www.d4science.eu)
  - iMarine (www.i-marine.eu)

- Available as a WebApp in the D4Science e-Infrastructure

http://www.d4science.org/