



The importance of the interoperability in Lifewatch

Nicola Fiore*, Daniel Fuentes**, Alberto Basset*,
Juan Miguel González-Aranda***

* Università del Salento

** LifeWatch ICT Team Spain

*** Spanish Ministry of Economy and Competitiveness



Biodiversity

Which actions to ensure long-term sustainability

Can we adapt to environmental change

What are the impacts of changes in climate, pollution and land/sea-use on biodiversity

How to manage multi-functional land/sea-scapes

Where are the thresholds in ecosystem structures and functions

How do changes affect the provision of ecosystem services

?

The image is a composite of four panels. The left panel shows an aerial view of a coastal landscape with a river, fields, and a beach. The top right panel shows a blue background with a DNA double helix and a fishing net. The bottom right panel shows a close-up of a mussel on a rocky shore. A large yellow question mark is centered between the left and right panels.

Biodiversity

Information extraction from noisy data
in complex systems

Network stability
and optimization

Transport effects;
Extreme network effects

Analysis and modeling
of interacting systems
Chaotic processes

Behaviour of individuals;
Irrational and chaotic factors;
Profits – costs interactions

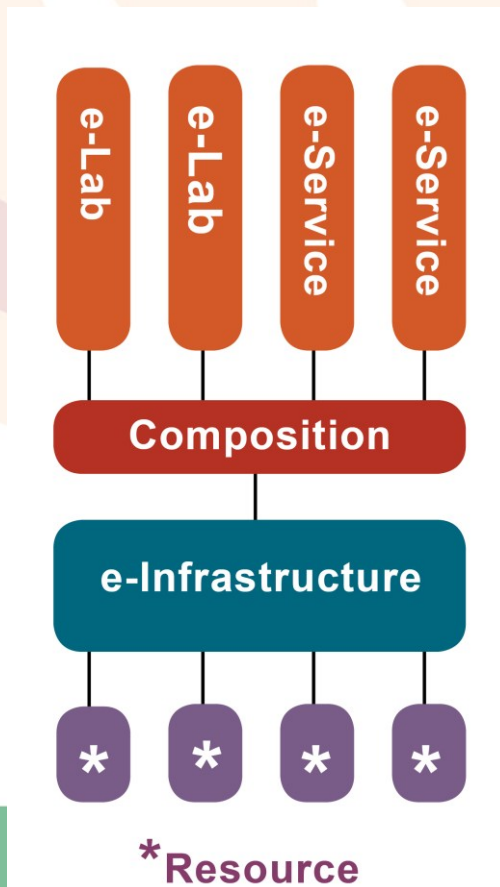
Scaling
effects

Modeling for
decision support

Interactions between microscopic (local)
and macroscopic (global) levels

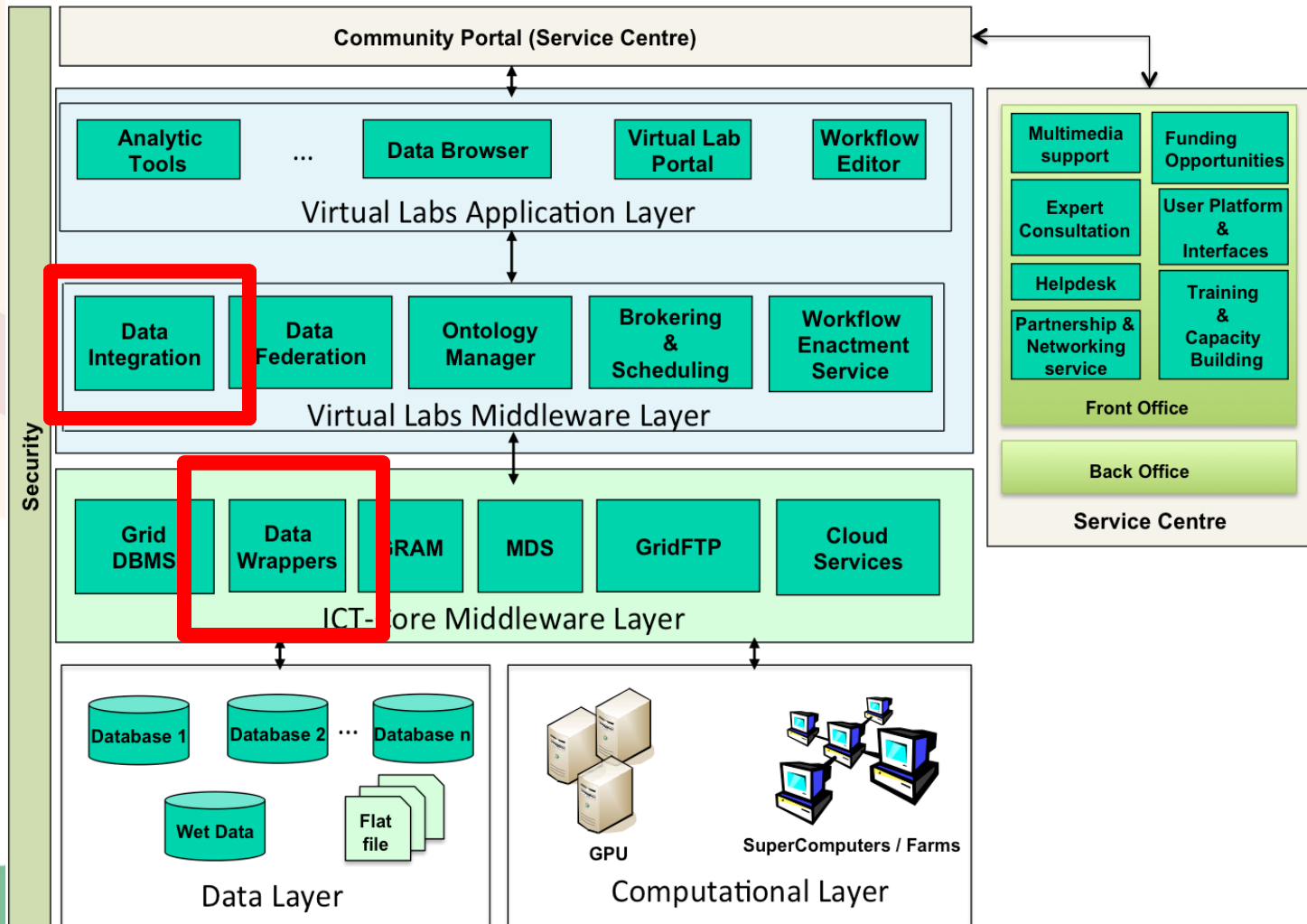
LifeWatch project

- LifeWatch is a community driven e-infrastructure



- User groups can create their own e-laboratories or e-services.
 - The e-laboratories are the '*community driven*' infrastructure, which promotes innovation.
 - Sharing data and algorithms scientists can address questions not otherwise accessible.
- * Italy – Service Center
 - * Spain – Statuory Seat
 - * The Netherlands - Centre IT Research

LifeWatch architecture

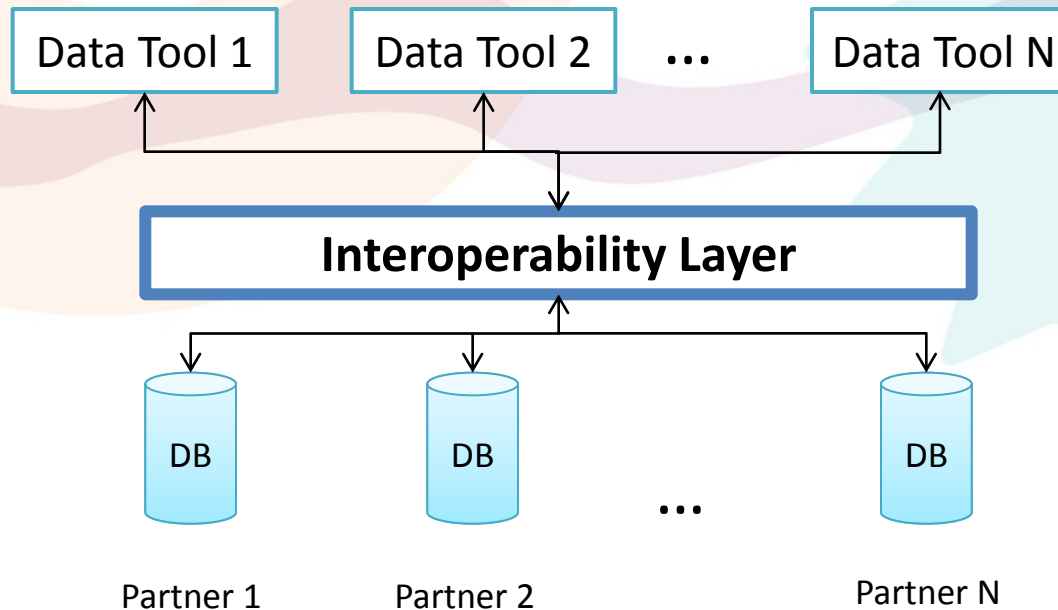


Interoperability in LifeWatch



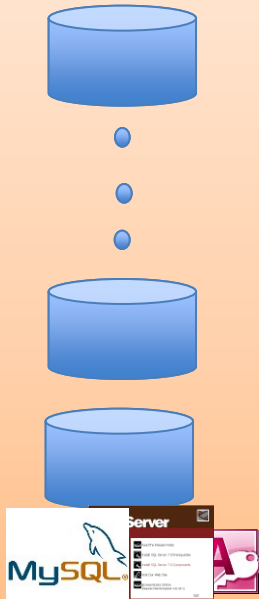
Interoperability in LifeWatch

- Lifewatch data tools (e.g. data portal) manage data coming from different datasources (with different data, physical location or schema).

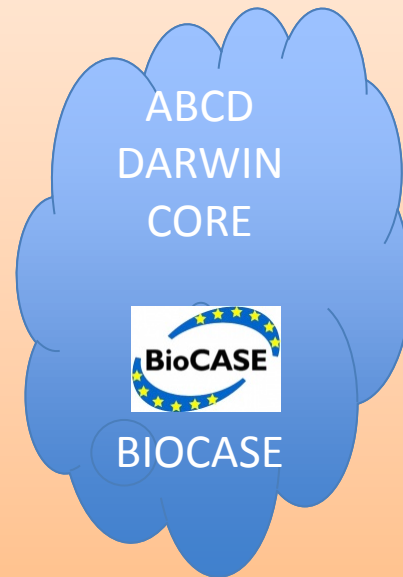


Interoperability in LifeWatch

Input DATA PROVIDERS



INTEROPERABILITY STANDARD/TOOLS



DATA CLEANING



www.faunaeur.org

www.fishbase.org

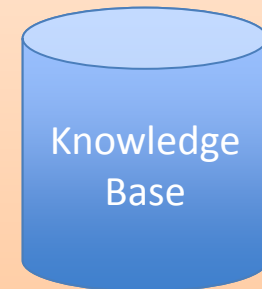
www.algaebase.org

www.actaplantarum.org

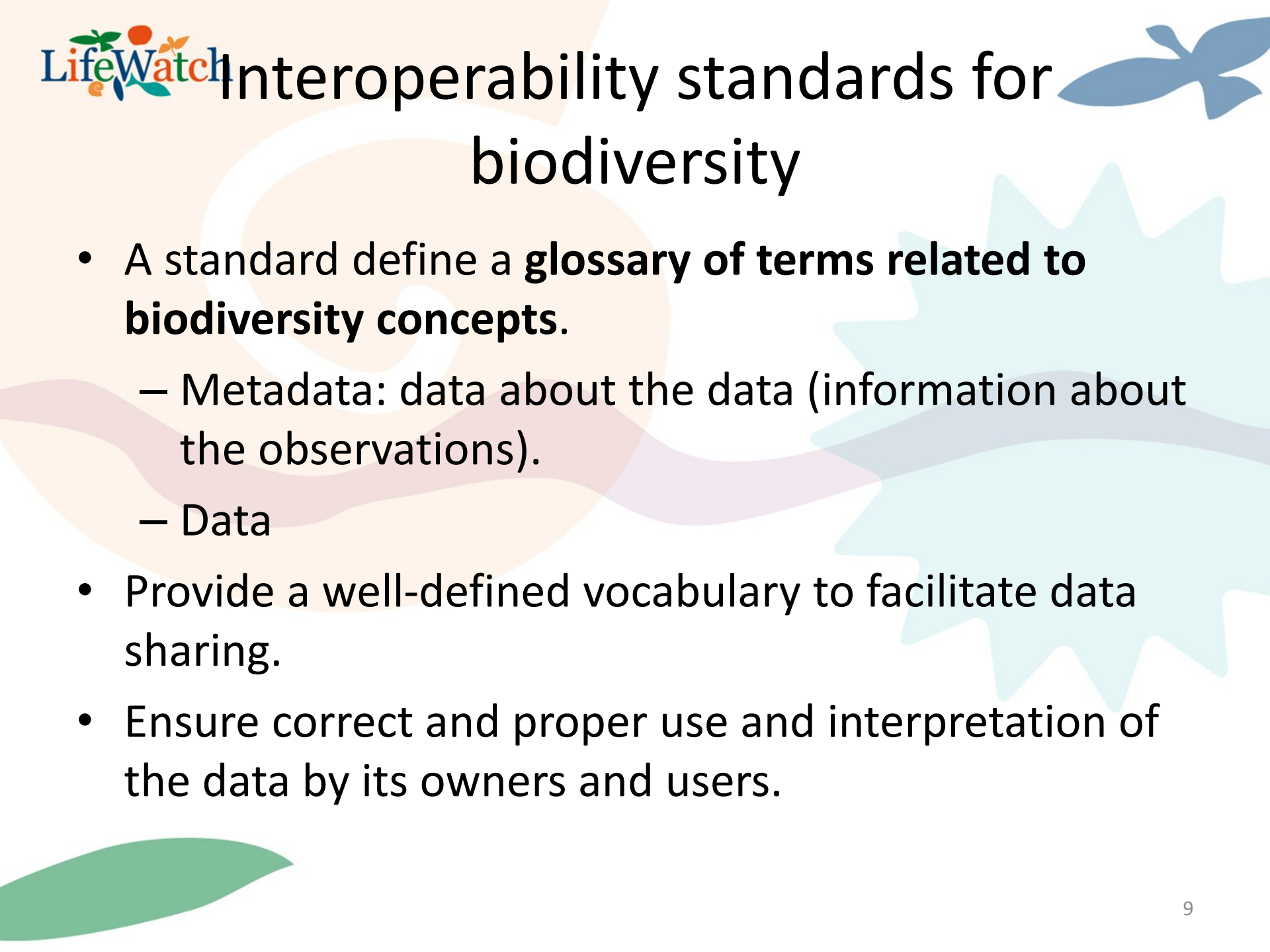
www.ittiofauna.org

www.eunis.org

Output KNOWLEGDE BASE



Interoperability standards for biodiversity

The background features several decorative elements: a large orange sun-like shape in the upper left, a blue bird silhouette in the upper right, a light blue starburst shape on the right side, and a green leaf shape in the bottom left corner.

- A standard define a **glossary of terms related to biodiversity concepts**.
 - Metadata: data about the data (information about the observations).
 - Data
- Provide a well-defined vocabulary to facilitate data sharing.
- Ensure correct and proper use and interpretation of the data by its owners and users.

Interoperability standards for biodiversity

Darwin Core v. 1.4 (2009)

- 90 concepts for spatiotemporal occurrences.
- Extensions for curatorial, geospatial and paleontology areas.
- Organizations: TDWG, GBIF.

ADVANTAGE: Simplicity, usability.

DISADVANTAGE: Weak description.

```

<metadata xml:lang="en">
  <dc:title xml:lang="en">Test</dc:title>
  <dc:type>http://purl.org/dc/dcmitype/Service</dc:type>
  <accesspoint>http://localhost/tapirlink/tapir.php/Test</accesspoint>
  <dc:description xml:lang="en">Test</dc:description>
  <dc:language>en</dc:language>
  <dc:subject xml:lang="en">Test</dc:subject>
  <dc:bibliographicCitation xml:lang="en">Test</dc:bibliographicCitation>
  <dc:rights xml:lang="en">Test</dc:rights>
  <dc:modified>2012-05-30T10:17:34</dc:modified>
  <dc:created>2012-05-30T10:08:54</dc:created>
  <relatedEntity>
    <role>data supplier</role>
    <role>technical host</role>
  </relatedEntity>
  <entity type="person">
    <identifier>Test</identifier>
    <name xml:lang="en">Test</name>
    <acronym>Test</acronym>
    <description xml:lang="en">Test</description>
    <address>Test</address>
  </entity>

```

Interoperability standards for biodiversity

ABCD v. 2.06 (2005).

- Preserved, living and data collections, DNA.
- 1200 concepts.
- Geosciences, DNA, Herbarium.
- TDWG, Universität Berlin.

ADVANTAGE: Detailed description.

DISADVANTAGE: Ambiguity.

```

<biocase:response xsi:schemaLocation="http://www.biocase
- <!--
  XML generated by BioCASE PyWrapper software v
-->
- <biocase:header>
  <biocase:version software="os">nt</biocase:version>
- <biocase:version software="python">
  2.6 (r26:66721, Oct 2 2008, 11:06:43) [MSC v.1500 6
</biocase:version>
  <biocase:version software="pywrapper">3.0</biocase:
  <biocase:version software="dbmod">MS SQL Server r
  <biocase:sendTime>2012-11-26T16:11:03.726000</bioc
  <biocase:source>PaperExplorer@localhost</biocase:so
  <biocase:destination>127.0.0.1</biocase:destination>
  <biocase:type>search</biocase:type>
</biocase:header>
- <biocase:content recordCount="10" recordDropped="0"
- <abcd:DataSets>
  - <abcd:DataSet>
    - <abcd:TechnicalContacts>
      - <abcd:TechnicalContact>
        <abcd:Name>LW Spain</abcd:Name>
      </abcd:TechnicalContact>

```

Interoperability standards for biodiversity



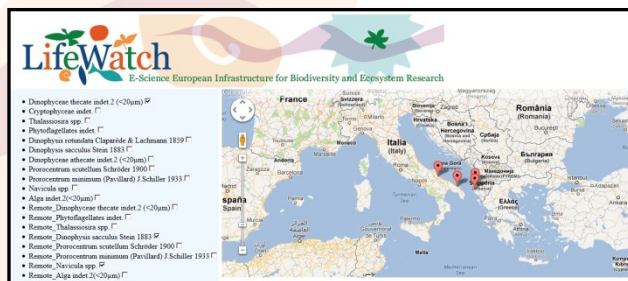
- **INSPIRE**

- Infrastructure for Spatial Information in Europe.
- Based on the spatial ISO standards from the ISO191XX series (so based on GML).
- Divided in 34 spatial data themes needed for environmental applications.
- Our conclusions about INSPIRE:
 - The specifications are not yet finished and developed tools are not enough matured.
 - Biodiversity occurrence data does not directly fall into a specific INSPIRE.
 - For now, we decide to wait before starting to use INSPIRE.

Demo

Show the location in a google maps of a set of species.

- Standards: Darwin Core, ABCD.
- Tools for the mapping: Tapir (Darwin Core) and Biocase (ABCD).



Interoperability Layer

<FullScientificNameString >
(ABCD)



Local Server

<ScientificName >
(Darwin Core)



Remote Server



Data portal using Biocase



LifeWatch **Query ABCD Data Porvider**

Limit max. number of results per database to
 Wait for not more than sec. for a database to respond (consider raising this timeout level if you get insufficient results).

- FullName
- NamedArea
- TitleCitation
- HigherTaxonName
- HigherTaxonRank
- AuthorTeamAndYear
- SpecieInformalNameString
- SpecieNameComments
- DetailCitation

Database Connection

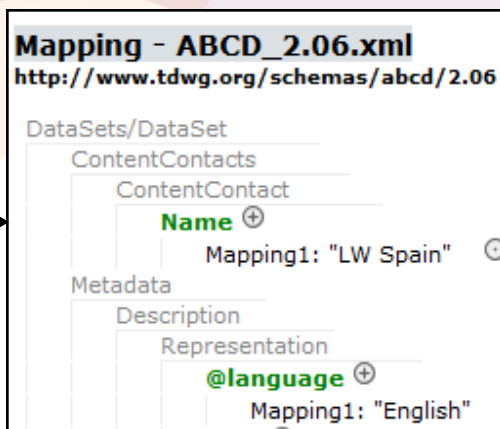
Status **OK**

DBMS Select the db_modu

Host The host of the data

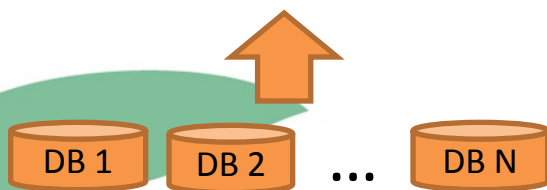
Database The name of the da

User



```

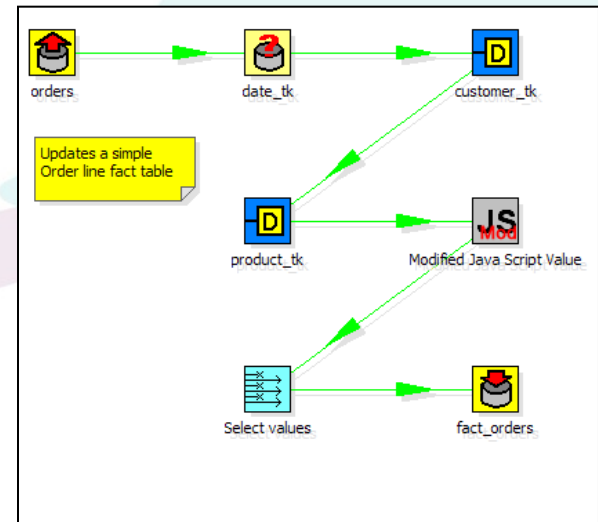
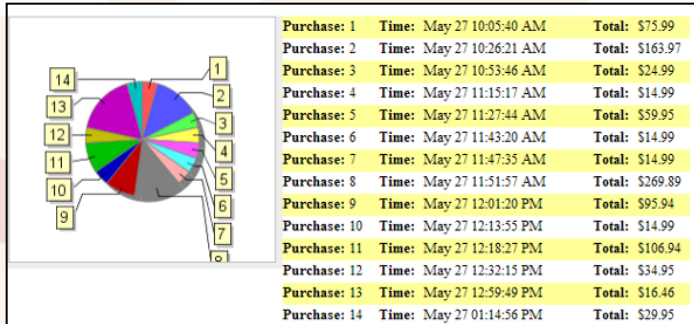
- <abcd:TechnicalContact>
  <abcd:Name>LW Spain</abcd:Name>
</abcd:TechnicalContact>
</abcd:TechnicalContacts>
- <abcd:ContentContacts>
- <abcd:ContentContact>
  <abcd:Name>LW Spain</abcd:Name>
</abcd:ContentContact>
</abcd:ContentContacts>
- <abcd:Metadata>
  
```





Pentaho

- Business intelligence tools for reporting, data mining, ETL, etc.



Pentaho Cube Designer <2>

Map Tables

1 2 3 4 5

ROOT QUERY
SELECT
FROM []
WHERE
GROUP BY
HAVING
ORDER BY

public Tables

- access
- access_log
- account_type_levels
- action_item
- action_item_log
- action_item_work
- action_item_work_notes
- action_item_work_selection
- action_list
- action_phase

ticket

- ticketid
- org_id
- contact_id
- problem
- entered
- enteredby
- modified
- modifiedby
- closed
- pri_code
- level_code
- department_code
- source_code
- cat_code
- subcat_code1
- subcat_code2
- subcat_code3
- assigned_to
- comment
- solution
- scode
- critical
- notified

lookup_ticket_status

- code
- description
- default_item
- level
- enabled
- entered
- modified

lookup_ticket_state

- code
- description
- default_item
- level
- enabled
- entered
- modified

access

- user_id
- username
- password
- contact_id
- role_id
- manager_id
- startofday
- endofday
- locale
- timezone
- last_in

Program Visualization Tools Help

WEKA
The University of Waikato

Applications

- Explorer
- Experimenter
- KnowledgeFlow
- Simple CLI

Waikato Environment for Knowledge Analysis
Version 3.6.2
(c) 1999 - 2010
The University of Waikato
Hamilton, New Zealand

Pentaho into Liferay



LifeWatch E-Science European Infrastructure for Biodiversity and Ecosystem Research

Home | LifeWatch History | Governance & Management >> | Participating Countries >> | LifeWatch in the Media | LifeWatch Service Centre | Communication Tools >> | Show Cases >>

LifeWatch is a European research infrastructure in development. The first services to users are planned for 2013.

Users may benefit from integrated access to a variety of data, analytical and modeling tools as served by a variety of collaborating initiatives. Another service is offered with data and tools in selected workflows for specific scientific communities. In addition, LifeWatch will provide opportunities to construct personalized "virtual labs", also allowing to enter new data and analytical tools.

New data will be shared with the data facilities cooperating with LifeWatch.

Download the LifeWatch brochure

LIFEWATCH SERVICE CENTER INAUGURATED

On 24 and June 2012, the LifeWatch Service Centre was inaugurated at its home base in Lecce, Italy. The Service Centre operates as the front office of LifeWatch with its scientific and other users, as well with the multitude of scientific organizations and projects currently or potentially cooperating with the infrastructure (read more) >

News

Workshop GARR - Computation and Distributed Storage
The second Workshop of the GARR-CSD is scheduled for 29-30 th Nov. [...]

Nordic LifeWatch
A pre-project funded by NordForsk. Joint Nordic efforts to [...]

The Exchange of Experience (EoE) Working Group of the CoPoRI project
The Exchange of Experience (EoE) Working Group of the CoPoRI proj [...]

pentaho open source business intelligence™

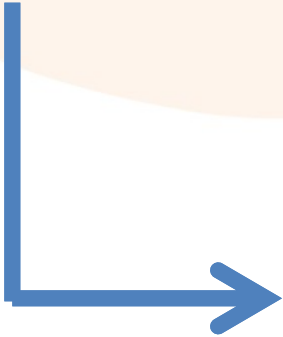
File View Tools Help

Browse

- Steel Wheels
 - Analysis
 - Charts
 - Dashboards
 - Reporting

Files

New Report | New Analysis View | Manage Content



LIFERAY Enterprise Open Source For Life

Home | Plugins | Pentaho | Add Page

Filters

Select filters to apply to other controls on this page

Select a Region
 Central Eastern Southern Western

Select one or more Departments
 Choose...
 Executive Management
 Finance
 Human Resource
 Marketing & Communication

Select one or more Position Titles
 Choose...
 CEO
 CMO
 IS
 EoE

Year

Update...

Actual Headcount - % Variance from Budget

CEO	SVP Partnerships	SVP Strategic Development	SVP WW Operations
5	-6	-5	-34

Headcount Costs

Category	Value	%
SVP WW Operations	476,000	(27%)
CEO	549,625	(31%)
SVP Strategic Development	383,242	(22%)
SVP Partnerships	367,415	(21%)

Headcount Data

Position	Actual	Budget	Variance
CEO	\$549,625	\$522,250	-\$27,375
Total	\$549,625	\$522,250	-\$27,375



Software Tools

- Other tools
 - GBIF Darwin Core Archive Validator.
 - GBIF Darwin Core Archive Assistant.
 - GBIF Integrated Publishing Toolkit (IPT).
 - Tapir*
 - Digir *

*Deprecated project.

Conclusions

- We will establish a way to exchange information considering different standards.
- We will establish the different steps to incorporate new datasources to the Lifewatch infrastructure.
- Developed tools do not need to consider the implementation details of the datasources.

Future work

- Manage the workflow of the data using Pentaho and Liferay.
- Design and develops discovery and reporting tools.
- Add new standards (i.e. INSPIRE directive) and cover new concepts (i.e. standard extensions).

The background is white with several large, stylized, overlapping shapes in shades of orange, pink, and light blue. A blue bird silhouette is in the top right, and a green leaf is in the bottom left.

Thanks!